

GLOBAL MAPPING OF EARTH-LIKE EXOPLANETS FROM SCATTERED LIGHT CURVES

HAJIME KAWAHARA^{1,2} AND YUKA FUJII²

ABSTRACT

Scattered lights from terrestrial exoplanets provide valuable information about the planetary surface. Applying the surface reconstruction method proposed by Fujii et al. (2010) to both diurnal and annual variations of the scattered light, we develop a reconstruction method of land distribution with both longitudinal and latitudinal resolutions. We find that one can recover a global map of an idealized Earth-like planet on the following assumptions: 1) cloudless, 2) a face-on circular orbit, 3) known surface types and their reflectance spectra 4) no atmospheric absorption, 5) known rotation rate 6) static map, and 7) no moon. Using the dependence of light curves on the planetary obliquity, we also show that the obliquity can be measured by adopting the χ^2 minimization or the extended information criterion. We demonstrate a feasibility of our methodology by applying it to a multi-band photometry of a cloudless model Earth with future space missions such as the occulting ozone observatory (O3). We conclude that future space missions can estimate both the surface distribution and the obliquity at least for cloudless Earth-like planets within 5 pc.

Subject headings: astrobiology – Earth – scattering – techniques: photometric

1. INTRODUCTION

Recent progress in observational techniques has revealed various physical properties of exoplanets beyond orbital parameters and planetary mass. Detections of the atmospheric components have been reported for several systems using spectroscopy at the planetary transit and secondary eclipse (e.g. Charbonneau et al. 2002; Vidal-Madjar et al. 2003, 2004; Tinetti et al. 2007; Swain et al. 2008, 2009). Interior compositions can be inferred from planetary mass and radius (e.g. Léger et al. 2009; Charbonneau et al. 2009). Constructions of thermal maps of the planetary atmosphere have been proposed by (Williams et al. 2006; Cowan & Agol 2008). A longitudinal thermal map of HD 189733b has been constructed by Knutson et al. (2007) based on the method proposed by Cowan & Agol (2008).

Nevertheless, an identification of planetary surface components still remains an ambitious challenge. One of the promising approaches is to use the scattered light of exoplanets through the direct imaging observation (e.g. Seager et al. 2000; Ford et al. 2001; Sudarsky et al. 2005). Ford et al. (2001) focused on the inhomogeneity of the Earth surface which causes diurnal variation of the scattered light. They computed the scattered light from a model Earth observed at a distance of 10 pc and showed that time variations of the scattered light in different photometric bands highly depend on the geological and biological features on the planetary surface such as ocean, land, and even vegetation. More detailed characterizations (including spectroscopy) of the scattered light of the Earth and its time variations are discussed both via Earthshine observations (e.g. Woolf et al. 2002; Arnold et al. 2002; Montañés-Rodríguez et al. 2006) and simulations (Tinetti et al. 2006a,b; Montañés-Rodríguez et al. 2006).

These studies have suggested a future possibility to investigate the surface of Earth-like exoplanets by the scattered light curves. We note that such time variations of the scattered light are also applicable to determine the rotation period from as shown by Pallé et al. (2008).

A variety of inversion techniques of the planetary surface from the scattered light curves have been proposed. Surprisingly, the first theoretical study of scattered light curves to make albedo maps has been performed at the beginning of the twentieth century although the author assume asteroid and satellites for the target (Russell 1906). Cowan et al. (2009) performed principal component analysis (PCA) on multi-band photometric data of the Earth observed by EPOXI (Extrasolar Planet Observation and Deep Impact Extended Investigation) mission, and extracted spectral features which roughly correspond to land and ocean. They also checked the time variation of these components and translated it to the longitudinal distribution of these components based on the formulation by Cowan & Agol (2008). Oakley & Cash (2009) paid attention to the gap of reflectivity between ocean and land, and reproduced a longitudinal map of land. Fujii et al. (2010) (hereafter F10) have developed a methodology to estimate the areas of ocean, soil, vegetation, and snow from multi-band photometry, and showed that the area of these components can be recovered from mock observations of a cloudless Earth.

Since these authors focused on diurnal variations in mapping the surface, the resultant maps have only longitudinal resolution with little information of the latitudinal distribution. One of the goals of the present paper is to develop a method to map inhomogeneous surfaces of exoplanets with both longitudinal and latitudinal resolutions using both diurnal and annual variations of the scattered lights.

We also consider the determination of the planetary obliquity from time variation of planetary light. The obliquity is an important property of Earth-like planets with its strong implications for climate, habitabil-

kawa_h@tmu.ac.jp

¹ Department of Physics, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan² Department of Physics, The University of Tokyo, Tokyo 113-0033, Japan

ity (*e.g.*, Williams & Kasting 1997; Williams & Pollard 2003) and planetary formation. N-body simulations of the final stage of terrestrial planet formation indicated that the distribution of the obliquity ζ is isotropic (Agnor et al. 1999; Chambers 2001; Kokubo & Ida 2007). Gaidos & Williams (2004) modeled the infrared light curves of exoplanets and showed how the obliquity affects an annual variation. Oakley & Cash (2009) also pointed out a possibility to determine the planet's obliquity. In this paper, we demonstrate that the obliquity is estimated simultaneously with the global map of the planet by analyzing the scattered light curves over its orbital period.

The rest of the paper is organized as follows. We first review the estimation method of the weighted area from multi-band photometry proposed by F10, and describe our methodology to reconstruct the planetary surface and the obliquity measurement in §2. Assuming a future satellite mission for the direct imaging of Earth-like planets, we apply our methodology to mock observations based on real data of the scattering properties of the Earth in §3. Finally we summarize our results in §4.

2. METHODS

Let us briefly summarize the reconstruction method of the planetary surface from scattered light curves by F10. Scattered light from the planetary surface is characterized by the bidirectional reflectance distribution function (BRDF), $f(\vartheta_0, \vartheta_1, \varphi; \lambda) [\text{str}^{-1}]$, where ϑ_0 , ϑ_1 , and φ are the incident zenith angle, the scattering zenith angle, and the angle between the incident and scattered light, respectively (Fig. [1] a). The BRDF is also a function of the wavelength λ . Using the BRDF at (ϕ, θ) fixed on the sphere (Fig. [1]b), the total scattered intensity at a given phase is provided by

$$I(\lambda) = F_*(\lambda) R_p^2 \int_{S_{\text{VI}}} f(\phi, \theta; \lambda) \cos \vartheta_0(\phi, \theta) \cos \vartheta_1(\phi, \theta) d\Omega, \quad (1)$$

where $F_*(\lambda)$ is the incident flux at wavelength λ , R_p is a planetary radius, S_{VI} is the surface area facing on the observer and illuminated by the host star (we call it the visible and illuminated area), and $d\Omega = \sin \theta d\theta d\phi$. The position on the surface (ϕ, θ) fully specifies ϑ_0 , ϑ_1 , and φ , and thus we write $f(\phi, \theta; \lambda) \equiv f(\vartheta_0(\phi, \theta), \vartheta_1(\phi, \theta), \varphi(\phi, \theta); \lambda)$.

In order to make a fitting model for the multi-band photometric data, F10 classified the planetary surface into four types: ocean, snow, soil and vegetation. Denoting the specific BRDF of each type k by $f_k(\vartheta_0, \vartheta_1, \varphi)$, one can write the local BRDF at position (ϕ, θ) as a summation of the specific BRDFs of the k -th type surface weighted by local cover fractions $m^{(k)}(\phi, \theta)$:

$$f(\phi, \theta; \lambda) = \sum_{k=1}^{N_C} f_k(\phi, \theta; \lambda) m^{(k)}(\phi, \theta), \quad (2)$$

where N_C is the number of surface types and $N_C = 4$ in this case. Assuming that the scattering is isotropic (Lambertian), they approximate the BRDF by wavelength-dependence of albedo $a_k(\lambda)$,

$$f_k(\phi, \theta; \lambda) = \frac{a_k(\lambda)}{\pi}. \quad (3)$$

Under the above assumption, equation (1) reduces to

$$I(\lambda) = F_*(\lambda) R_p^2 \sum_{k=1}^{N_C} \frac{a_k(\lambda)}{\pi} \times \int_{S_{\text{VI}}} m^{(k)}(\phi, \theta) \cos \vartheta_0(\phi, \theta) \cos \vartheta_1(\phi, \theta) d\Omega. \quad (4)$$

Assuming that they know spectra for each components, F10 developed a reconstruction method for the weighted area of the k -th component,

$$A_k \equiv R_p^2 \int_{S_{\text{VI}}} W(\phi, \theta) m^{(k)}(\phi, \theta) d\Omega$$

$$W(\phi, \theta) \equiv \frac{\cos \vartheta_0(\phi, \theta) \cos \vartheta_1(\phi, \theta)}{\int_{S_{\text{VI}}} \cos \vartheta_0(\phi, \theta) \cos \vartheta_1(\phi, \theta) d\Omega}, \quad (5)$$

where $W(\phi, \theta)$ is the weight function. The above definition reduces equation (4) to a set of linear discrete equation:

$$C \frac{I(\lambda_b)}{F_*(\lambda_b)} = \sum_{k=1}^{N_C} a_k(\lambda_b) A_k,$$

$$C \equiv \pi \int_{S_{\text{VI}}} \cos \vartheta_0(\phi, \theta) \cos \vartheta_1(\phi, \theta) d\Omega, \quad (6)$$

where C is a factor that depends on the phase angle, *i.e.*, the planetocentric angle between the host star and the observer, and λ_b is the center wavelength of the b -th band. Using a cloudless model Earth, they showed that the weighted area can be approximately recovered by solving equation (6) with an inequality condition $A_k > 0$, and the summation of $A_k(t)$ is approximately constant over time, $\sum A_k(t) \approx \text{const.} \approx R_p^2$. It means that one can roughly estimate R_p as well by their method. Therefore, we use the fractional area A_k/R_p^2 normalized by R_p^2 in what follows. F10 have considered the spin rotation of a planet only and recovered the weighted area during one day of the planet and the weighted area can be converted to longitudinal maps by the inversion method proposed by Cowan & Agol (2008) and Cowan et al. (2009).

2.1. An Inverse Problem of a Two-dimensional Map

In the present paper, we consider the orbital motion to recover a two-dimensional world map of the planetary surface. We assume a model planet on a face-on circular orbit, which is the most promising case for the planet mapping. We also assume that orbital period is $P_{\text{orb}} = 365$ [day] and spin rotational period is $P_{\text{spin}} = 24$ [hrs]. We denote the phases of the orbital motion and spin rotation by Θ and Φ (Fig. [1] c), respectively. Our mock observation is assumed to be performed over one year from $\Theta = \Theta_S$ to $\Theta = \Theta_S + 2\pi$. There are two unknown parameters in this geometry: the obliquity ζ and the initial orbital longitude Θ_S . We discuss the measurement of ζ and Θ_S in section 2.3.

For our conventional geometry, the weight function is

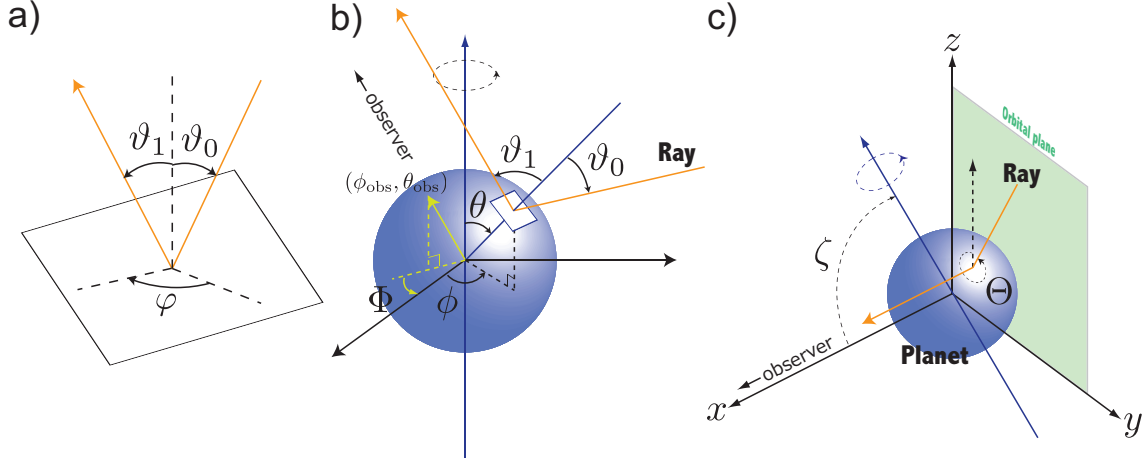


Figure 1. Schematic configurations of the planetary surface and system. Panel a) displays the definitions of the arguments of the BRDF $f(\vartheta_0, \vartheta_1, \varphi; \lambda)$: the incident zenith angle, ϑ_0 , the zenith angle of scattering, ϑ_1 , and the relative azimuthal angle between the incident and scattered light, φ . Panel b) shows the spherical coordinate fixed on the planetary surface: the complementary latitude, θ , and the longitude, ϕ . The spin rotation axis is indicated by $\theta = 0$. The spherical coordinate of the vector from the planetary center to observer is denoted by $(\phi_{\text{obs}}, \theta_{\text{obs}})$. Then, the spin rotation is described by $\Phi \equiv 2\pi - \phi_{\text{obs}}$. Panel c) illustrates the coordinate system to specify the phase of the planet. The Θ denotes the orbital longitude measured from the summer solstice (i.e. the angle between the incident ray and the projected vector of the spin rotation axis to the orbital plane). The obliquity ζ is the angle between the spin rotation axis and a normal vector of the orbital plane. In this paper, we assume that the line of sight is perpendicular to the orbital plane.

given by

$$\begin{aligned} W(\phi, \theta; \Theta; \Phi; \zeta) &= \frac{3}{2} \cos \vartheta_0 \cos \vartheta_1 \\ &= \frac{3}{2} [\cos \theta \cos \Theta \sin \zeta - \sin \theta \sin(\phi + \Phi) \sin \Theta \\ &\quad - \sin \theta \cos(\phi + \Phi) \cos \Theta \cos \zeta] \\ &\quad \times [\sin \theta \cos(\phi + \Phi) \sin \zeta + \cos \theta \cos \zeta]. \end{aligned} \quad (7)$$

We provide the explicit form of the visible and illuminated area, S_{VI} in Appendix A. Note that the weight function is normalized to unity by integrating over the sphere:

$$\int W(\phi, \theta; \Theta; \Phi; \zeta) d\Omega = 1. \quad (8)$$

Figure 2 shows the behavior of the weight function $W(\phi, \theta; \Theta; \Phi; \zeta)$ for cases of $\zeta = 0^\circ, 45^\circ$ and 90° . The weight function covers the region of $0^\circ < \theta < 90^\circ + \zeta$ and $0^\circ < \phi < 360^\circ$ as the planet rotates around the host star and its spin axis. As a result, the scattered light curves contain information about the land distribution up to $50(1 + \sin \zeta)\%$ of the total surface area in principle.

We consider how to deduce the local cover fraction $m^{(k)}(\phi, \theta)$ if a set of the weighted area throughout the orbital and spin rotational periods is given. In order to solve this problem, we apply one of the techniques of tomography, the linear inverse problem, to the data. We assume that the planetary surface consists entirely of land and ocean and do not consider the other components nor further decompose them into clouds, vegetation or soil in what follows. Therefore we use the land cover fraction $m(\phi, \theta)$ for the surface classification ($k = \text{land}$). The surface at (ϕ, θ) covered by land only (ocean only) is expressed by $m(\phi, \theta) = 1$ ($m(\phi, \theta) = 0$). We also denote the weighted area of land by $A(\Phi(t_i), \Theta(t_i); \zeta)$. In §3, we also consider the components of vegetation and soil using mock photometric data.

To begin with, let us discretize the observing time in N epochs and pixelize the planetary surface in M pixels. The weighted area recovered at the i -th epoch t_i is

written as

$$\begin{aligned} &A(\Phi(t_i), \Theta(t_i); \zeta) \\ &= \sum_{j | (s_j \cap S_{\text{VI}}) \neq \emptyset} \int_{s_j} W(\phi, \theta; \Theta(t_i); \Phi(t_i); \zeta) m(\phi, \theta) d\Omega \\ &= \sum_{j | (s_j \cap S_{\text{VI}}) \neq \emptyset} \langle m \rangle_{ij} \int_{s_j} W(\phi, \theta; \Theta(t_i); \Phi(t_i); \zeta) d\Omega, \end{aligned} \quad (9)$$

where s_j indicates the j -th pixel on the surface. The weighted cover fraction $\langle m \rangle_{ij}$ is defined as

$$\langle m \rangle_{ij} \equiv \frac{\int_{s_j} W(\phi, \theta; \Theta(t_i); \Phi(t_i); \zeta) m(\phi, \theta) d\Omega}{\int_{s_j} W(\phi, \theta; \Theta(t_i); \Phi(t_i); \zeta) d\Omega}. \quad (10)$$

Here, let us define the pixel-averaged cover fraction

$$\overline{m}_j \equiv \frac{\int_{s_j} m(\phi, \theta) d\Omega}{\int_{s_j} d\Omega}. \quad (11)$$

Under the assumption that the weighted cover fraction approximately equals to the pixel-averaged cover fraction:

$$\langle m \rangle_{ij} \approx \overline{m}_j, \quad (12)$$

we obtain

$$A(\Phi(t_i), \Theta(t_i); \zeta) = \sum_j \overline{m}_j \int_{s_j} W(\phi, \theta; \Theta(t_i); \Phi(t_i); \zeta) d\Omega. \quad (13)$$

We introduce the data vector of weighted area \mathbf{a} , the design matrix G and the model vector of cover fraction \mathbf{m} as

$$\begin{aligned} \mathbf{a} &\equiv \{a_i = A(\Phi(t_i), \Theta(t_i); \zeta) / \sigma_i \mid i = 1, 2, \dots, N\}, \\ G &\equiv \{G_{ij} = \int_{s_j} W(\phi, \theta; \Theta(t_i); \Phi(t_i); \zeta) d\Omega / \sigma_i \\ &\quad \mid i = 1, 2, \dots, N, j = 1, 2, \dots, M\}, \\ \mathbf{m} &\equiv \{\overline{m}_j \mid j = 1, 2, \dots, M\}, \end{aligned} \quad (14)$$

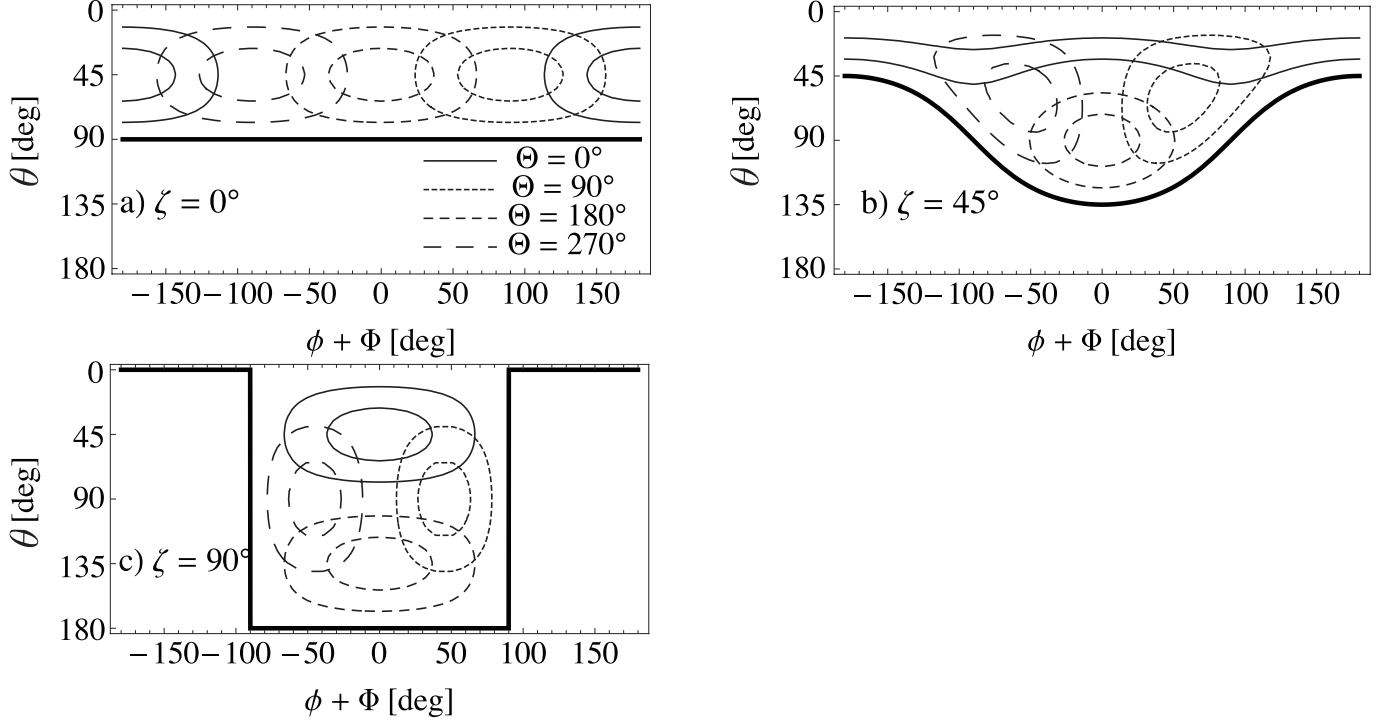


Figure 2. Geometry dependence of the weight function $W(\phi, \theta; \Theta; \Phi; \zeta)$ in the visible and illuminated area S_{VI} . Annual variations are indicated by different thin lines. Solid, large-dashed, middle-dashed and small-dashed lines indicate $\Theta = 0^\circ, 90^\circ, 180^\circ$, and 270° , respectively. Because the weight function solely depends on $\phi + \Phi$, in other words, longitude minus the phase of the spin rotation, we adopt $\phi + \Phi$ to x -axes instead of showing diurnal variations explicitly.

Thick lines indicate the boundary of the visible area, $\theta = \theta_V(\phi; \Phi; \zeta)$ in equation (A3). Each panel shows a different obliquity, $\zeta = 0^\circ$ (panel a), 45° (panel b), and 90° (panel c). Inner and outer contours indicate $W(\phi, \theta; \Theta; \Phi; \zeta) = 0.6$ and 0.3 , respectively.

where σ_i is the error of the weighted area. Then we can rewrite equation (13) as

$$\mathbf{a} = G \mathbf{m}. \quad (15)$$

The above equation can be solved for \mathbf{m} by minimizing χ^2 :

$$\chi^2 \equiv |\mathbf{a}_{\text{data}} - G \mathbf{m}|^2, \quad (16)$$

under inequality constraints:

$$0 \leq m_j \leq 1, \quad (17)$$

where \mathbf{a}_{data} is the observed weighted area with noise. We denote \mathbf{m} that minimizes equation (16) by \mathbf{m}_{est} .

2.2. Solving the Inverse Problem

We are now in a position to solve equation (16) with the condition (17) using an idealized data set. Since one can choose an arbitrary division for the surface modeling, we use the following pixelization:

$$\begin{aligned} s_j &= \{(\phi, \theta) | 1 - (1 + \sin \zeta) \frac{j_\theta}{M_\theta} < \cos \theta \leq \\ &1 - (1 + \sin \zeta) \frac{j_\theta - 1}{M_\theta}, \frac{2\pi}{M_\phi}(j_\phi - 1) \leq \phi < \frac{2\pi}{M_\phi} j_\phi\}, \\ j &= (j_\phi - 1)M_\theta + j_\theta \\ &\text{for } j_\phi = 1, 2, \dots, M_\phi \text{ and } j_\theta = 1, 2, \dots, M_\theta. \end{aligned} \quad (18)$$

The above pixelization has a constant solid angle for a given ζ and the number of pixels does not change for different ζ 's. We adopt $M_\phi = 9$ and $M_\theta = 23$ (see the right panel of Figure 3 for the $\zeta = 90^\circ$ case) and assume the numbers of epochs for diurnal variations $N_\Phi = 23$

and for annual variations $N_\Theta = 23$. In Appendix B, we discuss the dependence of model degeneracies on the pixelization.

We create the synthetic data of the weighted area using $1^\circ \times 1^\circ$ fixed land/water masks of the ISLSCP II (*input map*; left panel of Figure 3). Pixel averaged land cover fraction of this data (*reference map*) is shown in the right panel of Figure 3. We fix the obliquity $\zeta = 90^\circ$ and $\Theta(t_0) = \Theta_S = 0^\circ$ and compute the weighted area $A_{\text{input}}(\Phi, \Theta; \zeta = 90^\circ)$ from the input map (not from the reference map). We compute the data A_{data} by adding a Gaussian noise N_g to $A_{\text{input}}(\Phi, \Theta; \zeta = 90^\circ)$:

$$A_{\text{data}}(\Phi, \Theta) = A_{\text{input}}(\Phi, \Theta; \zeta = 90^\circ) + N_g. \quad (19)$$

We denote the standard deviation of the Gaussian random variable N_g by $\sigma \equiv \langle N_g^2 \rangle^{1/2}$.

The linear inverse problem (eq. [15]), or the equivalent least square problem (eq. [16]), under the inequality constraints (eq. [17]) can be solved with Bounded Variable Least Squares Solver (BVLS) developed by Lawson & Hanson (1974, 1995)¹. The BVLS is a generalization of the non-negative least square (NNLS) described in Lawson & Hanson (1974, 1995) and uses the QR decomposition (see also Menke 1989). The BVLS works for both overdetermined and underdetermined (ill-condition) problems. Even for ill-condition problem, the BVLS provides one of the (non-unique) solutions (Lawson & Hanson 1995).

By solving the least square problem (eq. [16]) with

¹ The original code of the BVLS is available through NETLIB (<http://www.netlib.org/lawson-hanson/index.html>)

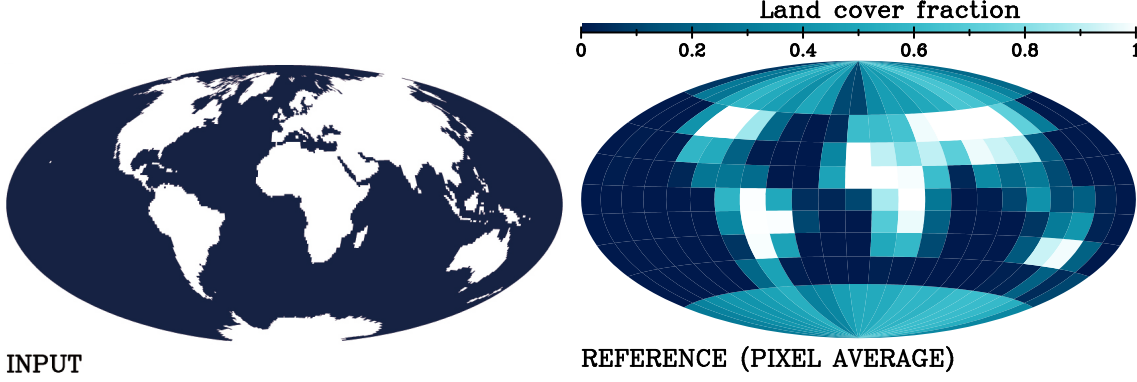


Figure 3. The input map (left) and the reference map (right). The left panel shows the land distribution of the Earth based on the data set of the ISLSCP II fixed land/water masks (<http://islsclp2.sesda.com>). Right panel indicates the pixel-averaged land cover fraction (\bar{m}_j) computed by averaging the left panel over pixels.

the BVLS algorithm, we obtain the recovered map \mathbf{m}_{est} . Figure 4 displays the two-dimensional images of the weighted area (left), the recovered maps (middle), and the prediction errors, $[A_{\text{data}}(\Phi(t_i), \Theta(t_i)) - G\mathbf{m}_{\text{est}}]/\sigma$ (right) for $\zeta_{\text{input}} = 90^\circ$. In this figure, we regard ζ and Θ_S as fixed parameters and use the input values $\zeta = \zeta_{\text{input}} = 90^\circ$ and $\Theta_S = \Theta_{S,\text{input}} = 0^\circ$ for equation (14). We consider four cases of the noise, $\sigma = 0, 0.01\langle A_{\text{input}} \rangle, 0.1\langle A_{\text{input}} \rangle$, and $0.3\langle A_{\text{input}} \rangle$, where $\langle A_{\text{input}} \rangle$ is the average of $A_{\text{input}}(\phi, \theta; \zeta = 90^\circ)$ and is approximately equal to 0.3 for the case of the Earth. As discussed in Appendix C, we confirmed that the estimated map is the unique solution under the inequality constraint. Comparing the recovered maps with the reference map (right panel in Fig. 3), we conclude that one can approximately recover the land distribution from the two-dimensional map of the weighted area.

We also solve the inverse problem in the case of the obliquity of the Earth, $\zeta_{\text{input}} = 23.45^\circ$ (Figure 5). Even for a small obliquity like the Earth, one can recover the land distribution of $50[1 + \sin(\zeta = 23.45^\circ)] \approx 70\%$ of the planetary surface. We conclude that our methodology can recover main features of the continents on the planetary surface even if the planet has small obliquity as is the case of the Earth.

2.3. Measurement of Obliquity

So far, we have assumed that the obliquity ζ and the initial phase Θ_S in the orbit are known a priori. Next we try to estimate the obliquity ζ by our method itself. Figure 6 indicates the obliquity dependence of images of the weighted area. We determine the best-fit value of ζ and Θ_S by minimizing χ^2 distance:

$$\chi^2(\zeta, \Theta_S) \equiv \left| \mathbf{a} - G(\zeta, \Theta_S) \mathbf{m}_{\text{est}}^{\zeta, \Theta_S} \right|^2$$

$$G(\zeta, \Theta_S) \equiv \{G_{ij}(\zeta, \Theta_S) = \int_{s_j} W(\phi, \theta; \Theta(t_i); \Phi(t_i); \zeta) d\Omega / \sigma_i$$

$$|i = 1, 2, \dots, N, j = 1, 2, \dots, M\},$$

$$\Theta(t_0) = \Theta_S, \quad (20)$$

where $\mathbf{m}_{\text{est}}^{\zeta, \Theta_S}$ is determined by minimizing χ^2 with given ζ and Θ_S using the BVLS fitting. Figure 7 displays $\chi^2(\zeta, \Theta_S)$ for $\zeta_{\text{input}} = 45^\circ$ and $\Theta_{S,\text{input}} = 180^\circ$. We use the *amoeba* routine (Press et al. 1992) to search for ζ and Θ_S , which minimizes $\chi^2(\zeta, \Theta_S)$. We estimate errors of ζ and Θ_S by the bootstrap resampling because

errors are originated not only from additional Gaussian noise but also from pixelization, linearization, and other systematics. After 10^3 iterations of the bootstrap, we obtain $\zeta_{\text{est}} = 46.3^\circ \pm 0.6^\circ$ and $\Theta_{S,\text{est}} = 181.8^\circ \pm 0.8^\circ$ for 10 % additional noise and $\zeta_{\text{est}} = 50.3^\circ \pm 2.3^\circ$ and $\Theta_{S,\text{est}} = 183.0^\circ \pm 2.3^\circ$ for 30 % additional noise. The estimated values of both obliquity and initial phase in orbit remarkably agree well with the input values.

3. APPLYING TO MOCK OBSERVATION OF MULTI-BAND PHOTOMETRY

Now, we apply our methodology to a mock multi-band observation of the cloudless Earth as a more realistic demonstration. We simulate scattered light curves of a cloudless Earth in the same manner as the simulation in F10. We use the solar spectrum as an incident flux on the model Earth. We assume an observer seeing the planet on a face-on and circular orbit at a distance of 5 pc (Case A) or 10 pc (Case B) from it. The scattered light from land is computed using the BRDF of actual land surface assigned by MODIS data-set (“snow-free gap-filled MODIS BRDF Model Parameters”). We adopt the data of April and ignore the seasonal variation of the BRDF at each point on the planetary surface. The scattered light from ocean is calculated with the BRDF model for wavy ocean described in Nakajima & Tanaka (1983). In our model, the snow cover regions around the poles are replaced by ocean. We also include the effect of Rayleigh scattering by atmosphere with single scattering approximation between the atmosphere and the underlying surface. Neither the effect of clouds nor the molecular absorption is, however, incorporated.

Our assumed observation parameters are listed in Table 1. As an observing system, we assume a satellite telescope with 1.1 m aperture. This assumption comes from the basic architecture proposed for the *occluding ozone observatory* (O3) (Kasdin et al. 2010), which is a satellite mission for direct imaging of exoplanets in UV/optical/near-IR bands shading the light from the host star by a 30 m external occulter. The scattered light is computed at the central wavelength of each band and multiplied by its band width. We consider four bands centered at the wavelengths listed in Table 2, which correspond to the bands of the MODIS land BRDF data. We assume $0.1 \mu\text{m}$ as the bandwidth.

We divide the orbital phase into 23 ($\Theta_i = 2\pi i/23$, $i = 0, 1, \dots, 22$), and the spin rotational phase into 23 ($\Phi_i = 2\pi i/23$, $i = 0, 1, \dots, 22$). We compute the scattered light

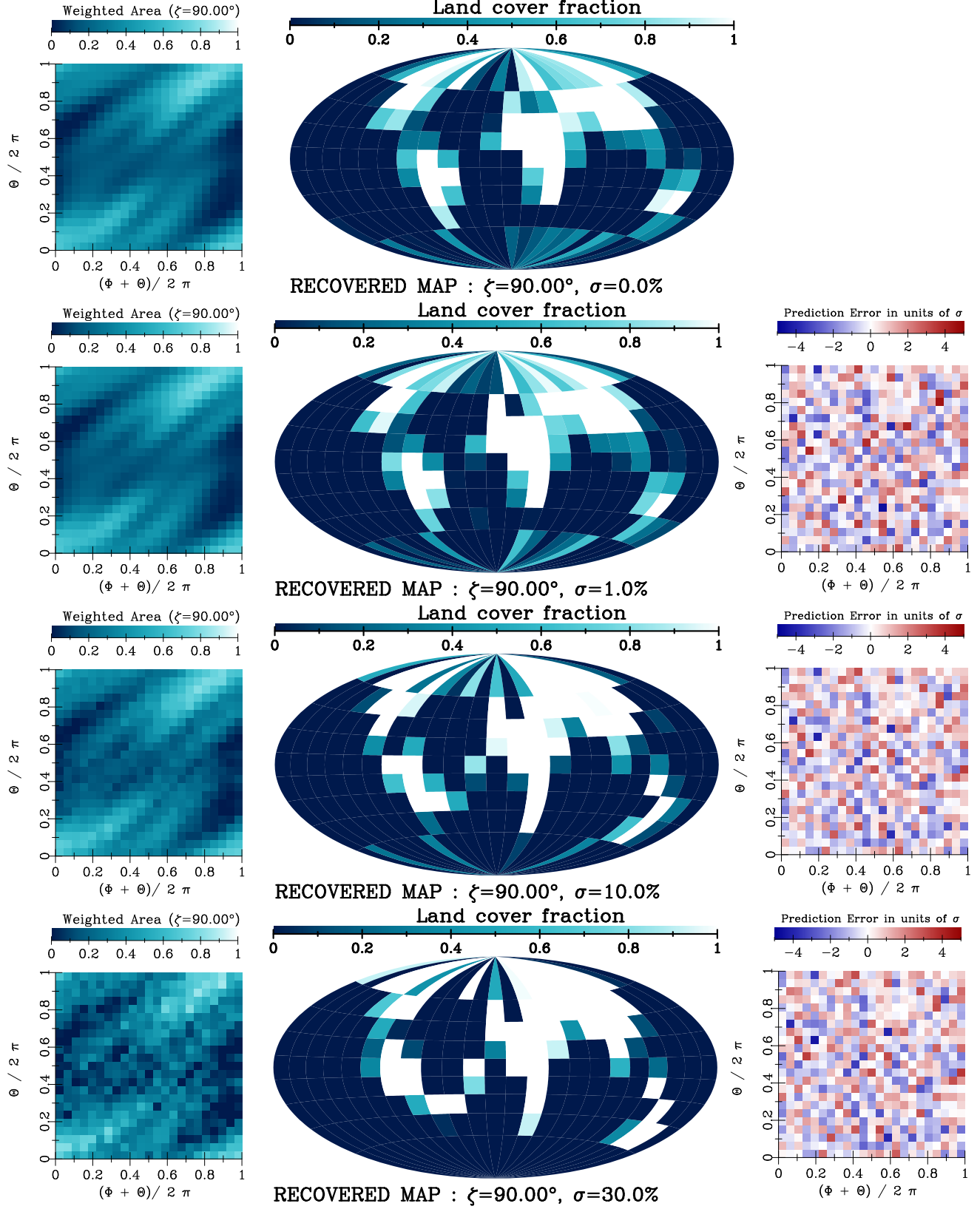


Figure 4. Two dimensional images of the weighted area $A(\phi, \theta)$ (left panels), recovered maps (middle panels) for $\zeta = 90^\circ$, and prediction errors, $[A_{\text{data}}(\Phi(t_i), \Theta(t_i)) - Gm_{\text{est}}]/\sigma$ for different noises, $\sigma = 0, 0.01\langle A_{\text{input}} \rangle, 0.1\langle A_{\text{input}} \rangle$, and $0.3\langle A_{\text{input}} \rangle$ from the top to the bottom panels.

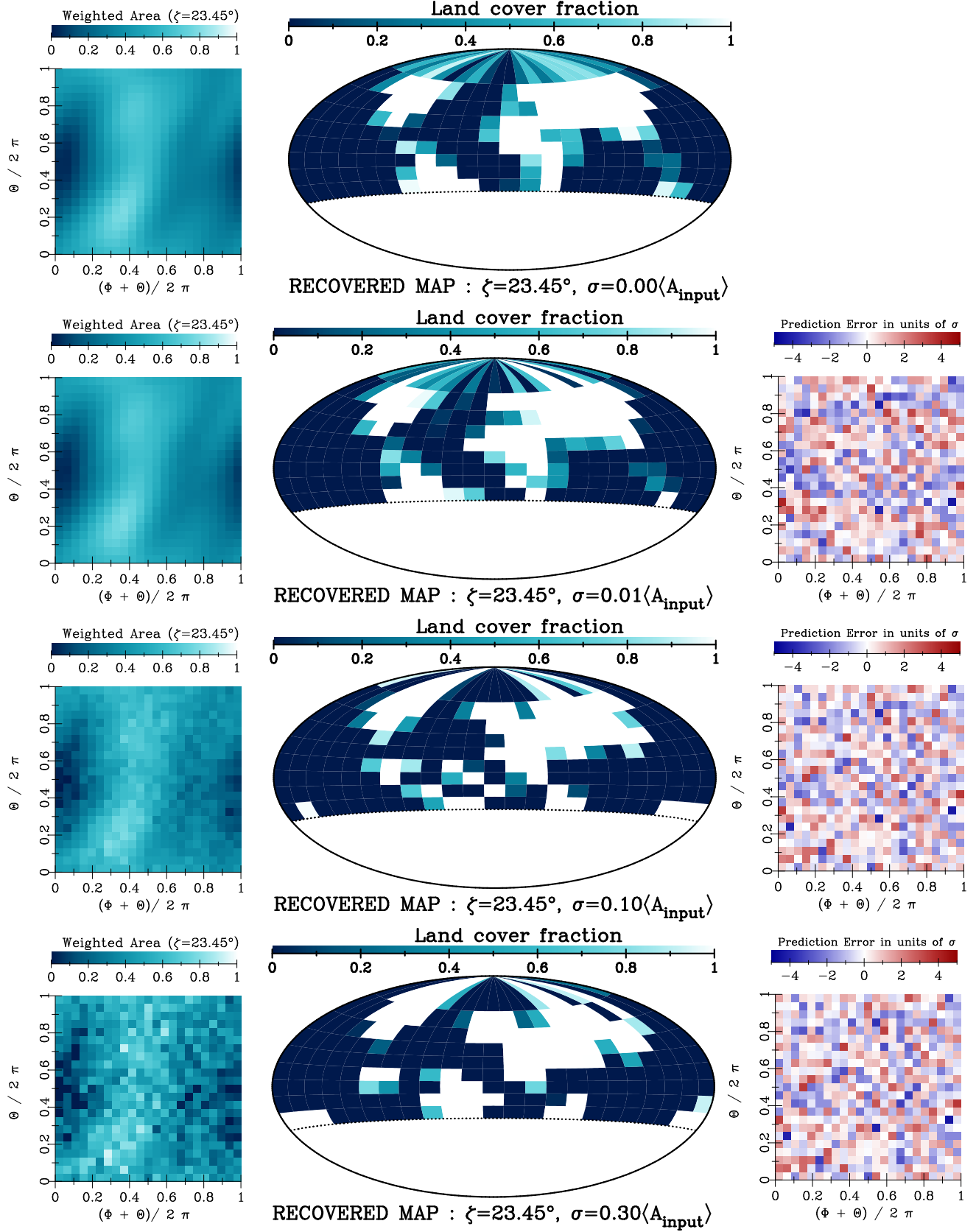


Figure 5. Two dimensional images of the weighted area $A(\phi, \theta)$ (left panels) and recovered maps (right panels) for $\zeta = 23.45^\circ$.

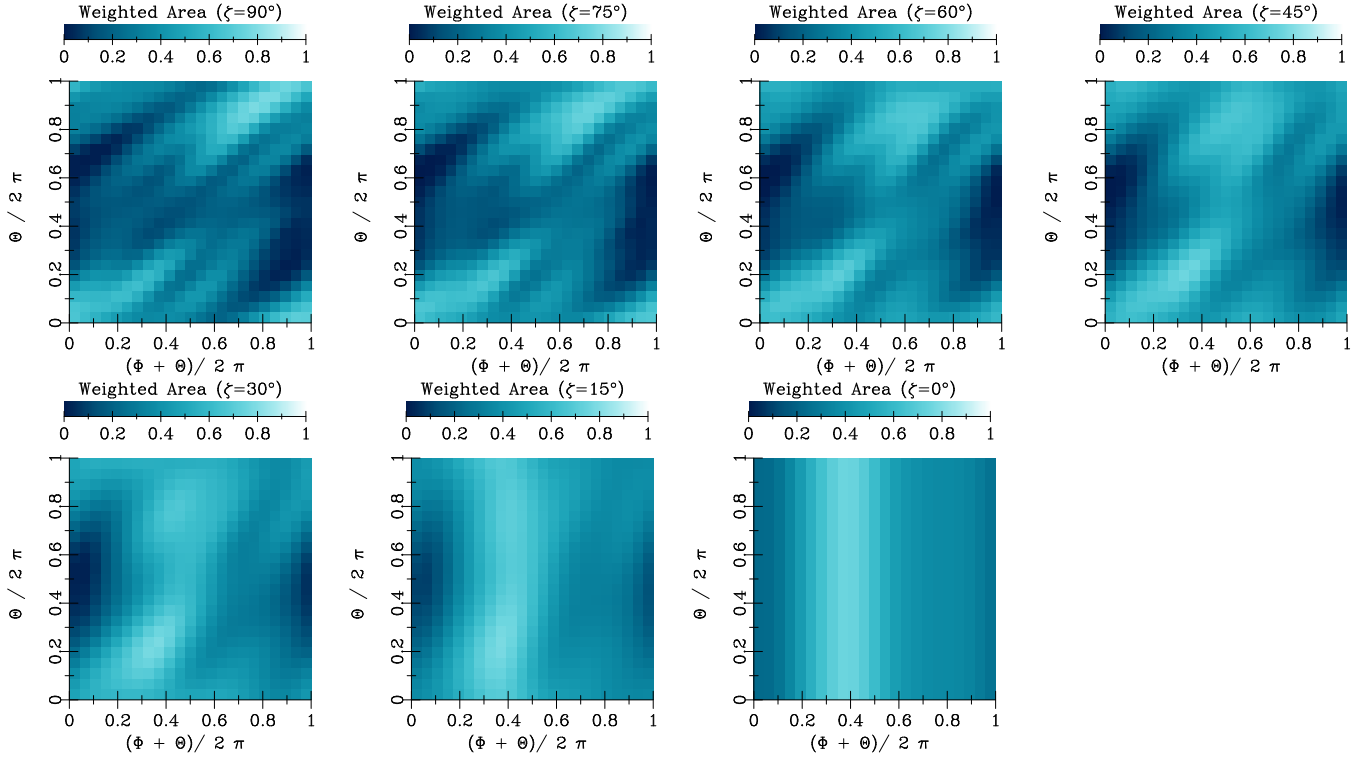


Figure 6. Obliquity dependence of two dimensional images of the weighted area.

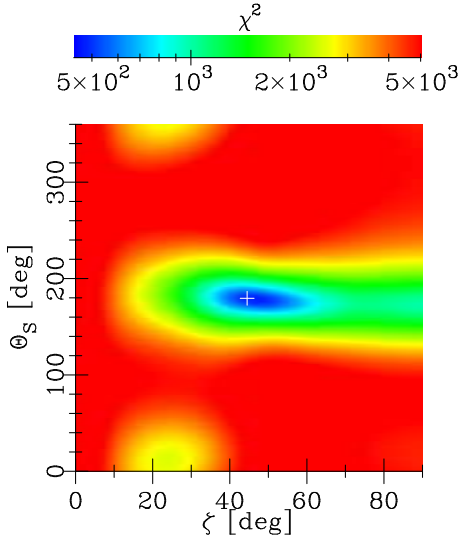


Figure 7. $\chi^2(\zeta, \Theta_S)$ for $\zeta_{\text{input}} = 45^\circ$ and $\Theta_{S,\text{input}} = 180^\circ$ (cross) with 10 % additional noise. The estimated obliquity and initial orbital longitude are $\zeta_{\text{est}} = 46.3^\circ \pm 0.6^\circ$ and $\Theta_{S,\text{est}} = 181.8^\circ \pm 0.8^\circ$, respectively.

curves in 4 bands with the exposure time $t_{\text{exp}} = 24/23$ hours throughout one year. Then, we stack data for 14 days centered at each orbital phase Θ_i , and fold the light curves according to its spin rotation period to obtain diurnal curves at the orbital phase, assuming that the spin rotational period is already known through periodogram analysis (Pallé et al. 2008). Thus, the resultant light curves have 23×23 data points for each band.

As for the noise, we consider the photon shot noise, the exzodiacal light, dark noise, and read noise. We ignore the leakage of the light from the host star. Thus, the

signal to noise ratio S/N is expressed as

$$S/N = \frac{S}{\sqrt{S + N_z + N_d + N_r}}, \quad (21)$$

where S , N_z , N_d , and N_r are the scattered light from the planet, the contribution from zodiacal light, dark noise, and read noise, respectively. Specifically, they are given

Table 1
Observation parameters

Symbol	Quantity	Value	Unit
t_{exp}	exposure time	24/23×3600	sec
n	folded days	14	days
D	diameter of telescope aperture	1.1	m
Ψ	sharpness	0.0433	
α	pixel scale	0.03125	arcsec/pixel
h	end-to-end efficiency	0.5	
v	dark rate	0.001	counts/sec
κ	read noise	2	$\sqrt{\text{counts}/\text{read}}$
QE	quantum efficiency	0.91	
Ω_z	zodiacal light in magnitude	23	mag/square arcsec
Z	zodiacal light	1	
F_0	zero flux	1.4×10^4 (band 1)	cts/cm ² /nm/s
		9.6×10^3 (band 2)	cts/cm ² /nm/s
		7.2×10^3 (band 3)	cts/cm ² /nm/s
		4.1×10^3 (band 4)	cts/cm ² /nm/s

Table 2
Band definition.

band	wavelength λ_b [μm]	band width $\Delta\lambda$ [μm]
1	0.469	0.1
2	0.555	0.1
3	0.645	0.1
4	0.8585	0.1

as

$$N_z \equiv hZF_0\alpha^2 10^{-\frac{\Omega_z}{2.5}} \frac{1}{\Psi} \pi \left(\frac{D}{2}\right)^2 t_{\text{exp}} n \Delta\lambda \quad (22)$$

$$N_d \equiv \frac{vt_{\text{exp}}n}{\Psi} \quad (23)$$

$$N_r \equiv \frac{\kappa^2 n}{\Psi}. \quad (24)$$

The parameters and their values we adopt are listed in Table 1 (see e.g., Savransky et al. 2010). The zero flux F_0 for calibration is calculated by linearly interpolating the spectrum of Vega given by Tables 1 and 2 in Colina et al. (1996), and shown in Table 3 in this paper. In order to quantitatively assess each noise level, we list the amplitudes of these noises per epoch in Table 3 together with the total average of the signal S in the case of $\zeta = 90^\circ$. The averaged signal-to-noise per epoch is also listed in Table 3. We note that the averaged signal per epoch is comparable with the noise level.

The obtained mock data $I(\lambda_b)$ at each epoch $(\Theta(t_i), \Phi(t_i))$ is decomposed into the weighted area of four surface types, (ocean, vegetation, soil, and snow) by solving equation (6) following F10. The weighted area of land is computed by summing up that of soil, vegetation, and snow, that is, $A(\Theta, \Phi) = A_{\text{soil}}(\Theta, \Phi) + A_{\text{vegetation}}(\Theta, \Phi) + A_{\text{snow}}(\Theta, \Phi)$. While F10 used nonlinear fitting in order to constrain A_k to be positive, we use the BVLS linear solver instead without upper constraints. The resultant $A(\Theta, \Phi)$ as a function of Θ and Φ is displayed in the left panels of Figure 8.

We now apply the reconstruction method described in §2 (see eqs [14][16][17]) to the set of obtained $A(\Theta, \Phi)$. For the errors σ_i in equation (14), we first compute the variance of the weighted area by 100 times resampling of

photon counts according to Poisson statistics, and we call such variance as $\sigma_{i,\text{res}}^2$. The $\sigma_{i,\text{res}}$, however, turned out not to be relevant for σ_i in equation (14) because it often results in zero or very small errors when $A_i = 0$ due to the boundary condition ($A_i \geq 0$). Therefore, we substitute an average value of errors $\sigma = \sum_{i=1}^N \sigma_{i,\text{res}}/N$ to equation (14) instead. For the case that all estimated errors of A_k are positive, the results with $\sigma_i = \sigma_{i,\text{res}}$ do not change significantly. In the fitting, we assume that the local cover fraction of each surface type remains unchanged and thus ignore the seasonal variations of vegetation, soil, and snow. The center and right rows of Figure 8 display recovered maps, and prediction errors for $\zeta_{\text{input}} = 90^\circ$ of the case A and B. The major features of the surface can be seen in the resultant maps. The averaged prediction errors $|A_{\text{data}}(\Phi(t_i), \Theta(t_i)) - Gm_{\text{est}}|/N$ are 0.53 $\langle A_{\text{data}} \rangle$ (case A) and 0.85 $\langle A_{\text{data}} \rangle$ (case B), respectively.

The land and ocean distributions have been considered so far. Next we create color composition maps of vegetation and soil on the land and ocean recovered map (Figure 9). Adopting the same method of the land recovery to the weighted areas of soil ($A_{\text{soil}}(t_i)$) and vegetation ($A_{\text{vegetation}}(t_i)$), we derive the soil and vegetation distributions and blend them by yellow (soil) and green (vegetation) over Figure 8, in other words, the land (white) and ocean (blue) distribution. Even using the weighted areas of soil and vegetation separately for the recovery, the resulting vegetation and soil maps almost distribute over the land that recovered with the summation of the weighted areas ($A_{\text{land}}(t_i) = A_{\text{soil}}(t_i) + A_{\text{vegetation}}(t_i) + A_{\text{snow}}(t_i)$). In this figure, snow is ignored because the contribution of it is significantly small. The concentrations of vegetation at South America and soil at middle Africa seen in the case A indeed correspond to the Amazon forest and the Sahara desert, respectively.

Now, we try to measure the obliquity for the mock observation. Figures 10, 11, 12, and 13 display the light curves of 4 bands with $\Theta_{S,\text{input}} = 180^\circ$, corresponding to different obliquities, $\zeta_{\text{input}} = 90^\circ, 60^\circ, 45^\circ$, and 30° , respectively. The noise assuming in these figures are computed in the “Case A”. Performing the χ^2 fitting described in §2.3, we estimate the best-fit obliquities and the initial orbital longitude. The predicted curves with the best-fit values are shown by solid lines in Figures 10 -

Table 3
Assumed signal and noises in 4 bands.

band b	1	2	3	4
averaged signal \bar{S} (case A) [10^1 cts/epoch]	40.9	40.3	36.9	40.9
averaged signal \bar{S} (case B) [10^1 cts/epoch]	10.2	10.1	9.2	10.2
exozodi N_z [10^3 cts/epoch]	12.4	9.2	7.0	4.0
dark noise N_d [10^3 cts/epoch]		1.2		
read noise N_r [10^3 cts/epoch]		1.3		
averaged S/N (case A) [1/epoch] *	3.3	3.6	3.7	4.9
averaged S/N (case B) [1/epoch] *	0.83	0.93	0.94	1.3

* The averaged S/N are defined by $\bar{S}/\sqrt{\bar{S} + N_z + N_d + N_r}$

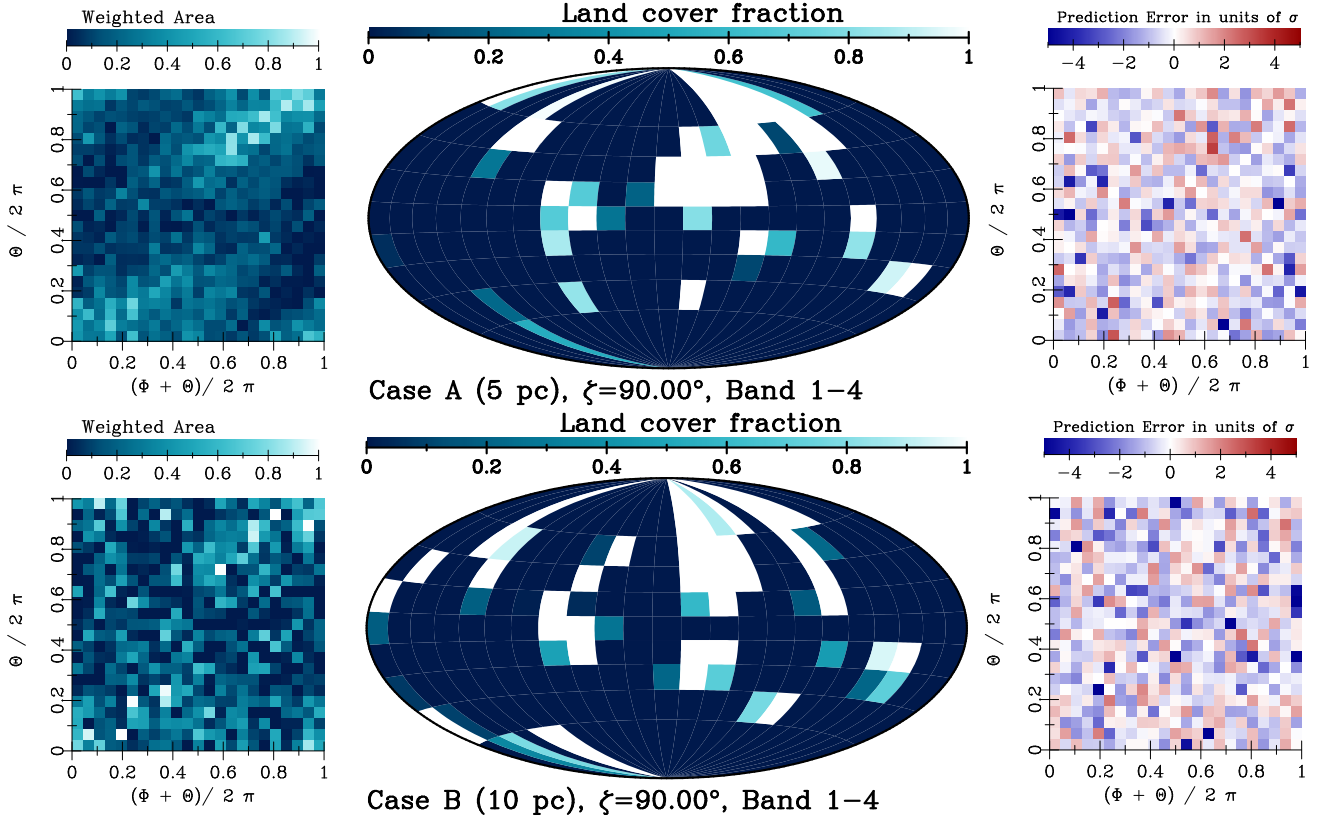


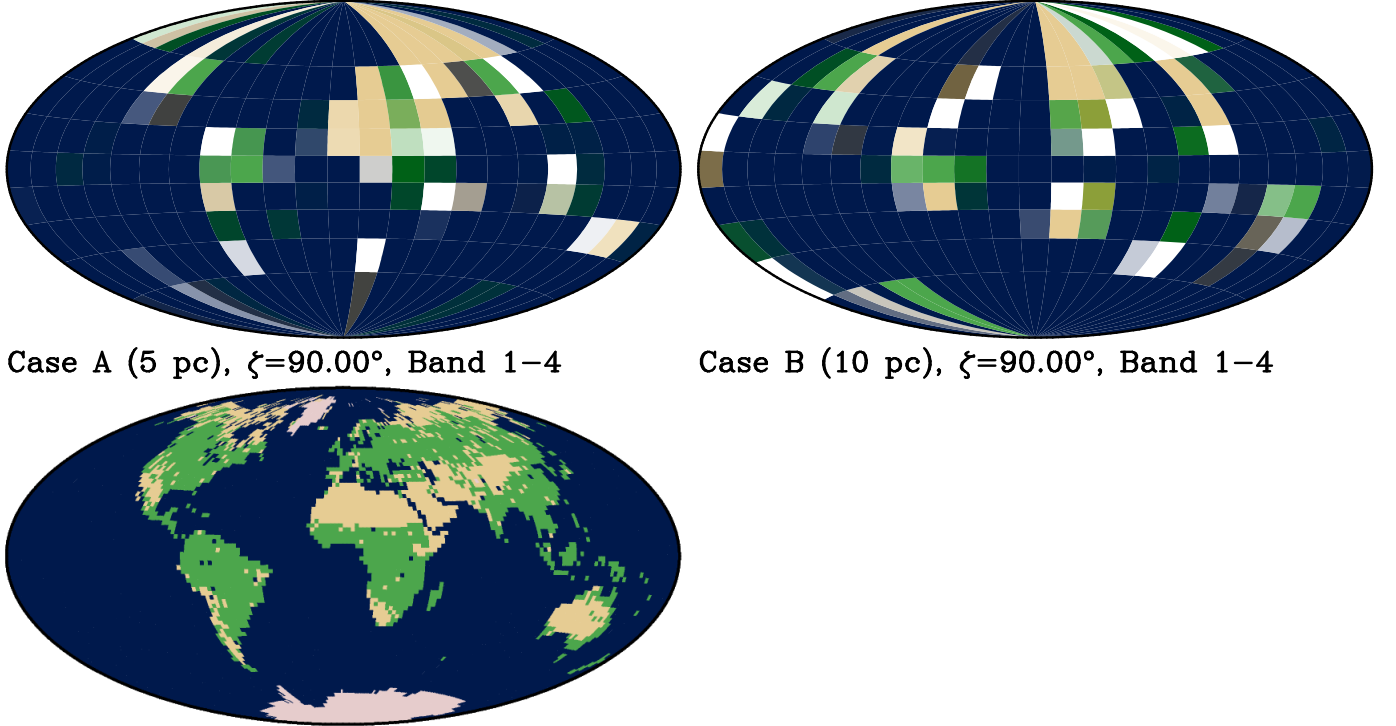
Figure 8. Simulated weighted area from mock photometric observation (left), its recovered map (middle), and prediction error (right). Top and bottom panels indicate the case A and B, respectively. We note that the snow component such as the South Pole in the input data is replaced to the ocean.

13. The input and the best-fit values and the reduced χ^2 are listed in Table 4. We compute the χ^2 for the land distribution (Eq. [20]). Therefore, the degree of freedom is $23 \times 23 - 23 \times 9 = 322$. It is likely that the inequality constraints and the other systematics such as the Lambert assumption induce the relatively high reduced χ^2 values (~ 2). One can see both diurnal and annual variations of light curves in these figures. The larger variations are seen in band 4 ($0.8585 \mu\text{m}$) due to large albedos of the land component (soil and vegetation; see Figure 7 in F10). There is “pinching” of the light curve at $\Theta_S = 0^\circ$ ($i \sim 264$) for $\zeta_{\text{input}} = \pi/4$ because the maximum of the weight function is located at the planet’s pole. In this period, the physical position of the maximum stays at the pole for the daily motion. As a result, there are little variations in the period. However the “pinching” seen at $\Theta_S \sim 180^\circ$ ($i \sim 0$) for $\zeta_{\text{input}} = 90^\circ$ has a different origin. In this period, the weight function moves around

the Southern Hemisphere, that is, ocean mainly.

The χ^2 maps for the four input obliquities are displayed in Figure 14 and the best-fit obliquities and initial longitudes are listed in Table 4. For the case A, the χ^2 fitting provides the estimated values close to the input ones within 10° . Figure 15 shows the difference of the predicted light curves with the best-fit obliquity $\zeta_{\text{est}} = 44.4$ and that shifted $\pm 3.3^\circ$ ($\sim 1\sigma$ level estimated by the bootstrap resampling), $\zeta = 47.7^\circ$ and 41.1° .

The results for the case B are also listed in Table 4. Figure 16 displays the χ^2 map for the case B. The estimated values for $\zeta_{\text{input}} = 30^\circ$ and 60° tend to deviate from the input one. As shown in the bottom right panel, the result for $\zeta_{\text{input}} = 30^\circ$ are significantly biased. It is likely that the systematics affect the obliquity estimates for this noise level. We also perform the other method called the extended information criterion in Appendix D. Although the tendency that the estimated value is biased

Case A (5 pc), $\zeta=90.00^\circ$, Band 1–4Case B (10 pc), $\zeta=90.00^\circ$, Band 1–4

CLASSIFICATION BASED ON IGBP

Figure 9. Recovered color maps of soil and vegetation from the mock observations. Soil and vegetation are blended by yellow and green on the land recovered maps shown in the middle panels of Figure 8. Top left and right panels indicate the case A and case B, respectively. Snow is replaced to ocean in this simulation. Bottom panel is a classification map of ocean (blue), snow (pink), vegetation (green), and soil (yellow) based on the IGBP classification. Our classification of soil and vegetation is listed in Table 2 of F10.

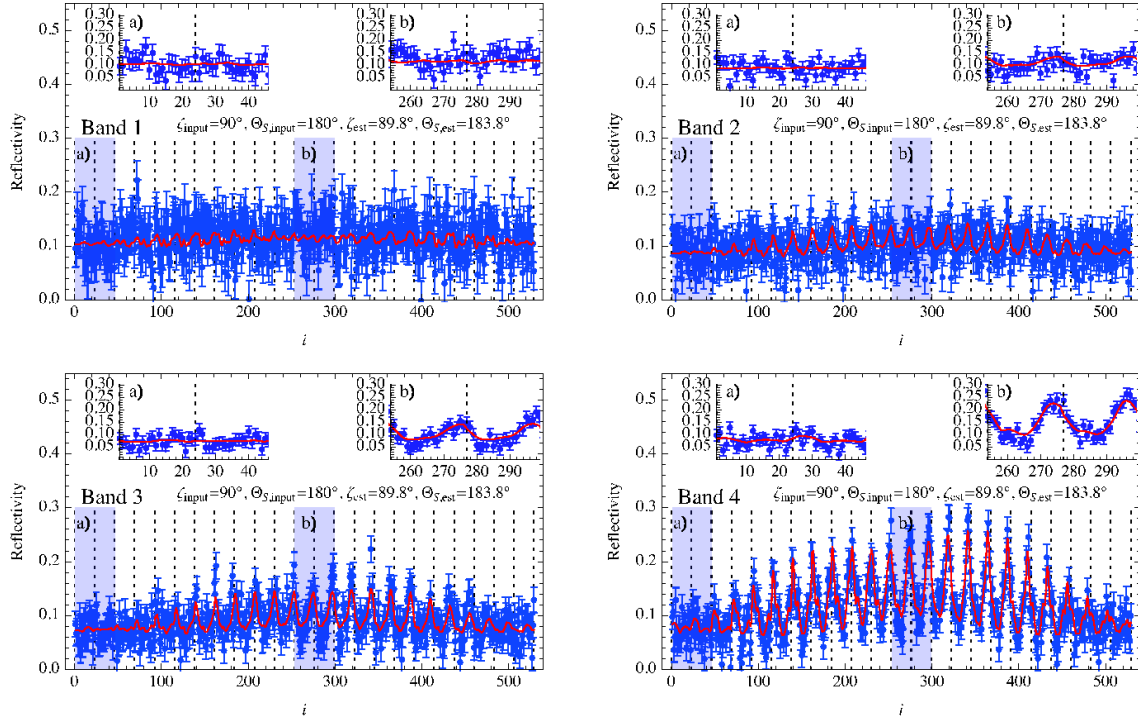


Figure 10. The mock light curves in units of reflectivity for $\zeta_{\text{input}} = 90^\circ$. Each panel indicates a different band: top left, top right, bottom left, and bottom right panels corresponds to bands 1 ($0.469\mu\text{m}$), 2 ($0.555\mu\text{m}$), 3 ($0.645\mu\text{m}$) and 4 ($0.8585\mu\text{m}$). Dashed vertical lines divide each epoch Φ , Θ stack in 14 days observations. There are 23 data points in one epoch and 23 epochs in whole data set. Therefore, there are $23 = 529$ data points in whole data set ($i = 1, 2, \dots, 529$). Wider versions of two shaded regions labeled in a) and b) are inserted in the mini panels. The predicted curves with the best-fit obliquities $\zeta_{\text{est}} = 89.8^\circ$ and initial longitude $\Theta_{S,\text{est}} = 183.8^\circ$ are drawn by solid lines. The observational noises are computed on the assumption of case A.

to $\zeta = 90^\circ$ vanishes, the best-fit value does not improve significantly.

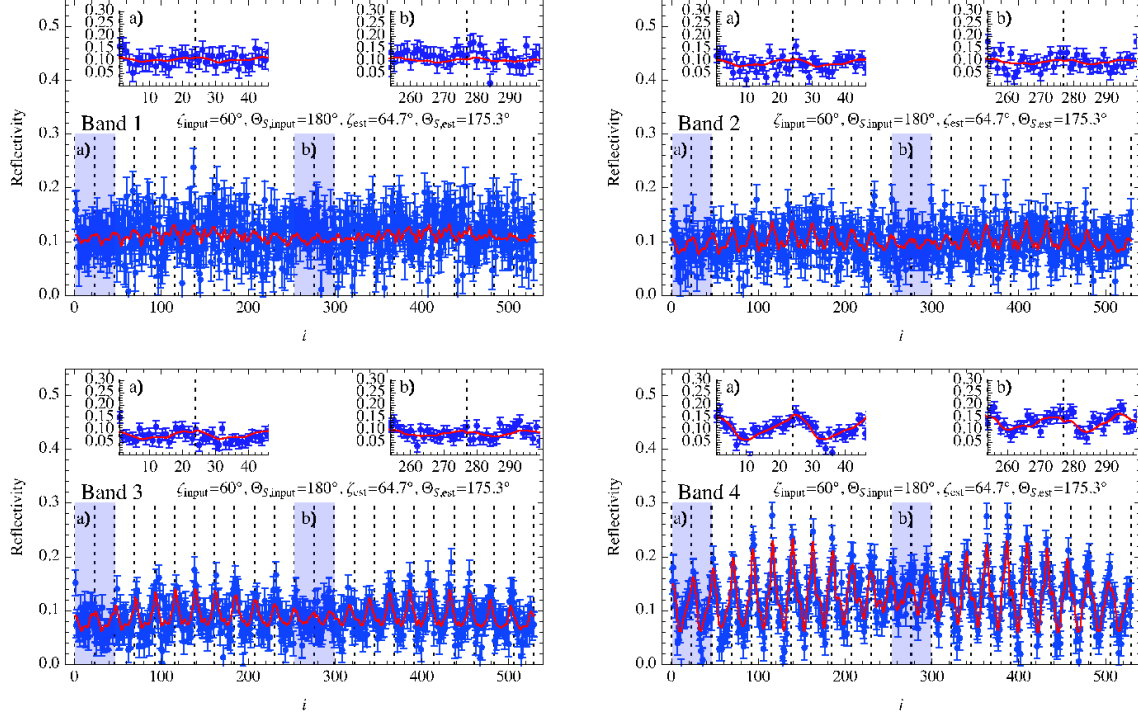


Figure 11. Same as Figure 10 for $\zeta_{\text{input}} = 60^\circ$.

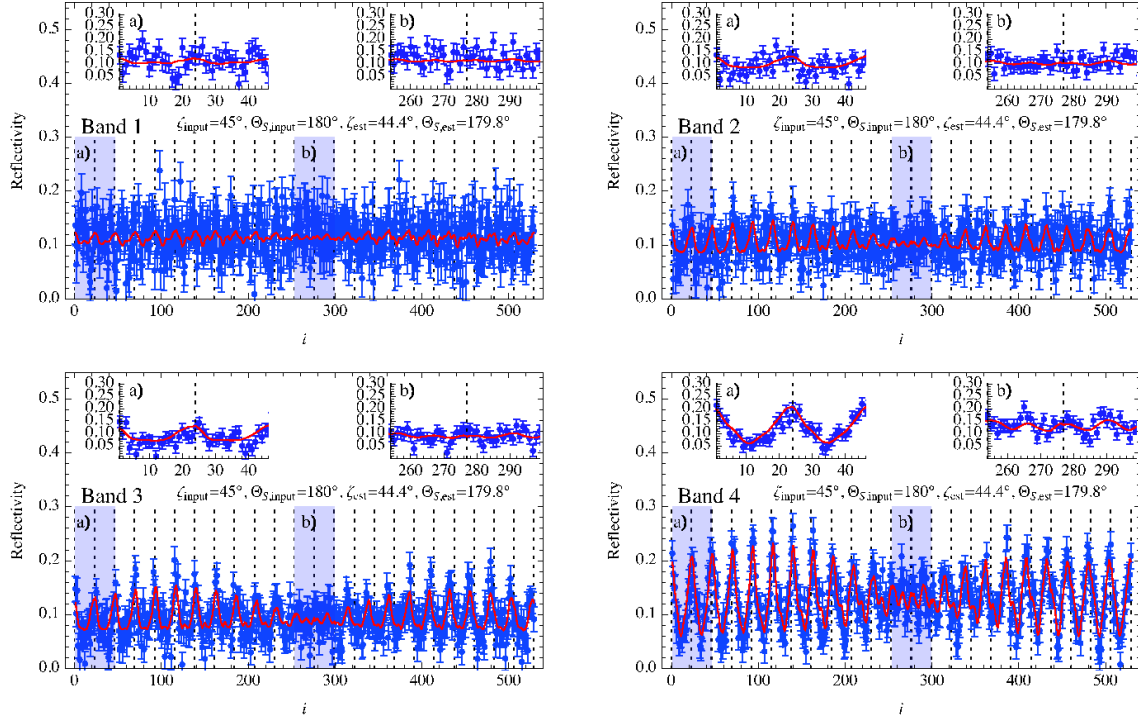


Figure 12. Same as Figure 10 for $\zeta_{\text{input}} = 45^\circ$.

In short summary, the planetary obliquity can be estimated well for the case A, while, for the case B that has larger noises, it is marginal to derive reliable values of the estimated obliquity by our method.

4. SUMMARY AND DISCUSSION

We have developed the reconstruction method of the two-dimensional planetary surface via diurnal and annual

variation of the scattered light. Applying the method to the mock photometric data, we have demonstrated that our method works for the mock Earth model, while this model has a lot of simplifying assumptions as follows: 1) cloudless, 2) a face-on circular orbit, 3) known reflectance spectra 4) no atmospheric absorption, 5) known rotation rate 6) static map, and 7) no moon. We also found that

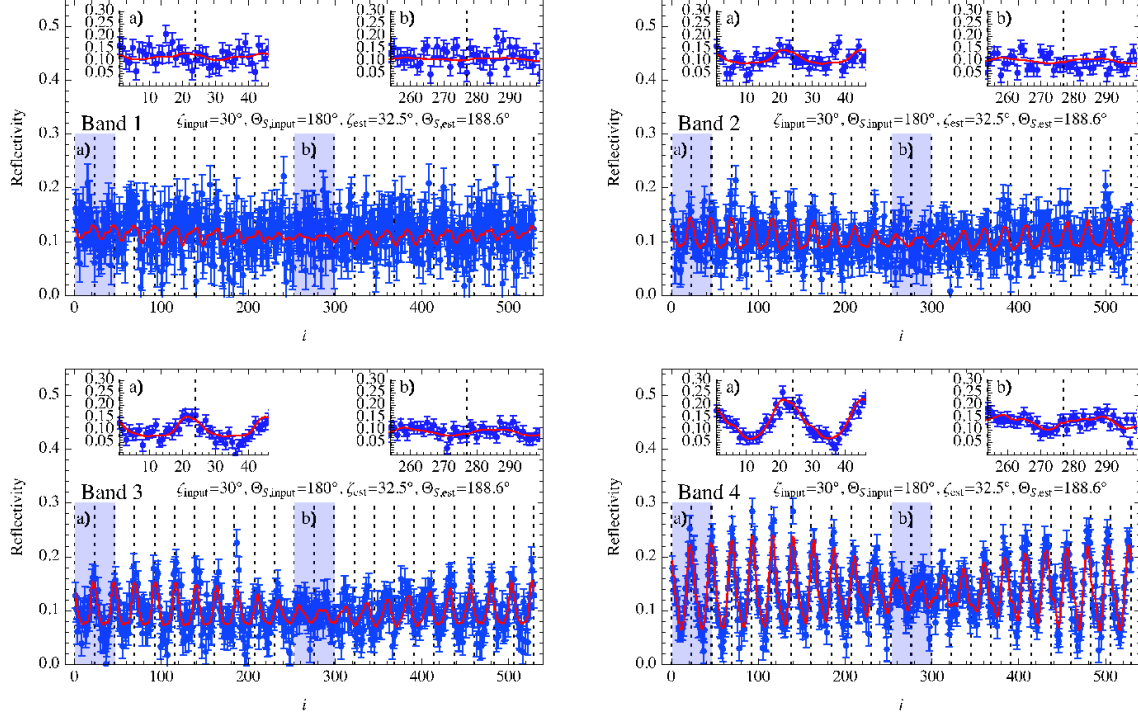


Figure 13. Same as Figure 10 for $\zeta_{\text{input}} = 30^\circ$.

Table 4
Estimated obliquity and the initial orbital longitude by χ^2 minimization and the EIC.

Case	ζ_{input}	$\Theta_{S,\text{input}}$	the χ^2 fit		χ^2	χ^2/dof	EIC	
			ζ_{est}	$\Theta_{S,\text{est}}$			ζ_{est}	$\Theta_{S,\text{est}}$
A	90°	180°	$89.8^\circ \pm 1.2^\circ$	$183.8 \pm 2.2^\circ$	650.3	2.02	89.0°	184.5°
A	60°	180°	$64.7^\circ \pm 4.8^\circ$	$175.3 \pm 3.5^\circ$	661.2	2.05	65.3°	176.5°
A	45°	180°	$44.4^\circ \pm 3.3^\circ$	$179.8 \pm 4.1^\circ$	731.1	2.27	44.5°	177.5°
A	30°	180°	$32.5^\circ \pm 3.0^\circ$	$188.6 \pm 5.6^\circ$	677.5	2.10	30.7°	190.5°
B	90°	180°	$89.8^\circ \pm 3.5^\circ$	$190.8^\circ \pm 17.5^\circ$	596.1	1.85	88.1°	187.5°
B	60°	180°	$79.7^\circ \pm 8.1^\circ$	$171.8^\circ \pm 24.8^\circ$	568.5	1.77	48.5°	155.6°
B	45°	180°	$38.2^\circ \pm 21.6^\circ$	$231.8^\circ \pm 43.6^\circ$	609.5	1.89	26.7°	132.6°
B	30°	180°	$89.2^\circ \pm 17.1^\circ$	$210.3^\circ \pm 32.7^\circ$	714.4	2.23	7.9°	138.6°

the planetary obliquity can be estimated by this method. With our method, future satellite missions such as the occulting ozone observatory (Kasdin et al. 2010) might provide “a global map” of Earth-like exoplanets. While only the terrestrial planets have been considered in this paper, our method might be applicable to any planets with an inhomogeneous surface, including Jupiter-like exoplanets.

In this paper, we ignored the effect of clouds and expected that clouds affect the estimation as like a statistical noise because of relatively short time variation of clouds. The PCA performed by Cowan et al. (2009) is one of promising approaches because they could separate the land and ocean compositions even though they used the EPOXI data that contains the cloud effect. The effect of clouds is discussed elsewhere (Fujii et al. in preparation). We also assumed a face-on circular orbit in this

paper. This assumption might be too severe for practical applications. We will generalize our method in the next paper.

We are deeply grateful to Yasushi Suto, Atsushi Taruya, and Edwin L. Turner for helpful and insightful discussion. We are thankful to Dmitry Savransky, Jeremy N. Kasdin, and David N. Spergel, who enlightened us about the observing instruments designed for future satellite missions. We also thank the referee, Nicolas Cowan for a lot of constructive comments. HK is supported by a JSPS (Japan Society for Promotion of Science) Grant-in-Aid for science fellows. This work is also supported by Grant-in-Aid for Scientific research from JSPS and from the Japanese Ministry of Education, Culture, Sports, Science and Technology (Nos. 20-10466 and 22-5467), and by the JSPS Core-to-Core Program “International Research Network for Dark Energy”.

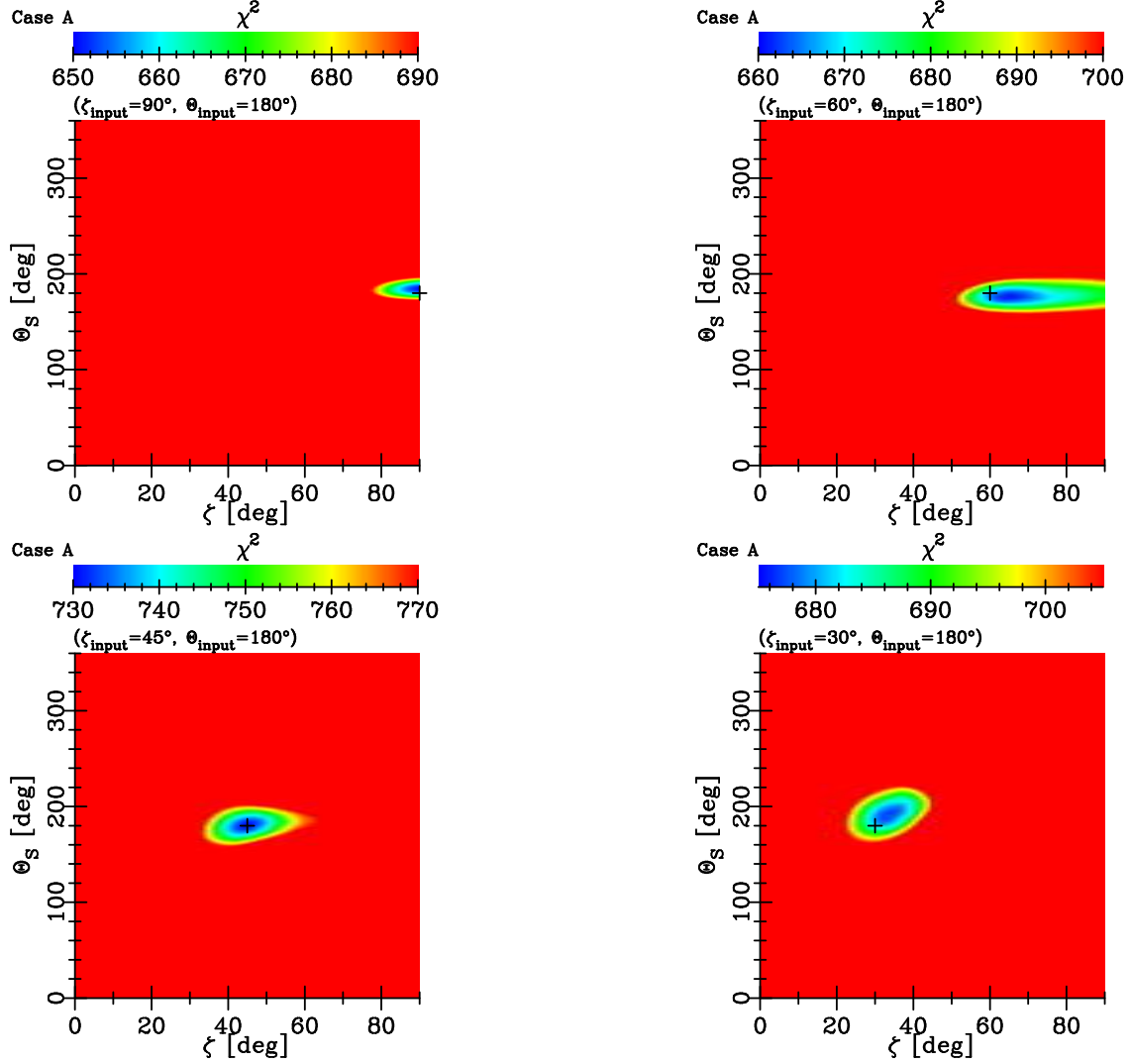


Figure 14. The ζ and Θ_S dependence of $\chi^2(\zeta, \Theta_S)$ of the mock observational data set (case A) for $\Theta_{S,\text{input}} = 180^\circ$ and the different input obliquities, $\zeta_{\text{input}} = 90^\circ$ (top left), $\zeta_{\text{input}} = 60^\circ$ (top right), $\zeta_{\text{input}} = 45^\circ$ (bottom left), $\zeta_{\text{input}} = 30^\circ$ (bottom right). The input values are indicated by cross. The estimated obliquity and initial orbital longitude are listed in Table 4.

APPENDIX

A. VISIBLE AND ILLUMINATED AREA

The visible and illuminated area is expressed as

$$S_{\text{VI}} = \begin{cases} \{(\phi, \theta) | 0 \leq \theta \leq \min[\theta_V(\phi; \Phi; \zeta), \theta_{I,+}(\phi; \Theta; \Phi; \zeta)]\} \\ \quad \text{for } 0 < \Theta < \pi/2 \text{ or } 3\pi/2 \leq \Theta < \pi, \\ \{(\phi, \theta) | \theta_{I,-}(\phi; \Theta; \Phi; \zeta) \leq \theta \leq \theta_V(\phi; \Phi; \zeta)\} \\ \quad \text{for } \pi/2 \leq \Theta < 3\pi/2. \end{cases} \quad (\text{A1})$$

The boundary lines of visible area $\theta_V(\phi; \Phi; \zeta)$ and illuminated area $\theta_{I,+}(\phi; \Theta; \Phi; \zeta)$, and $\theta_{I,-}(\phi; \Theta; \Phi; \zeta)$ are obtained by solving the following equation.

$$\begin{aligned} & W(\phi, \theta; \Theta; \Phi; \zeta) \\ &= (\sin \theta \cos(\phi + \Phi) \sin \zeta + \cos \theta \cos \zeta) \\ &\times (\cos \theta \cos \Theta \sin \zeta - \sin \theta \sin(\phi + \Phi) \sin \Theta - \sin \theta \cos(\phi + \Phi) \cos \Theta \cos \zeta) \\ &= 0. \end{aligned} \quad (\text{A2})$$

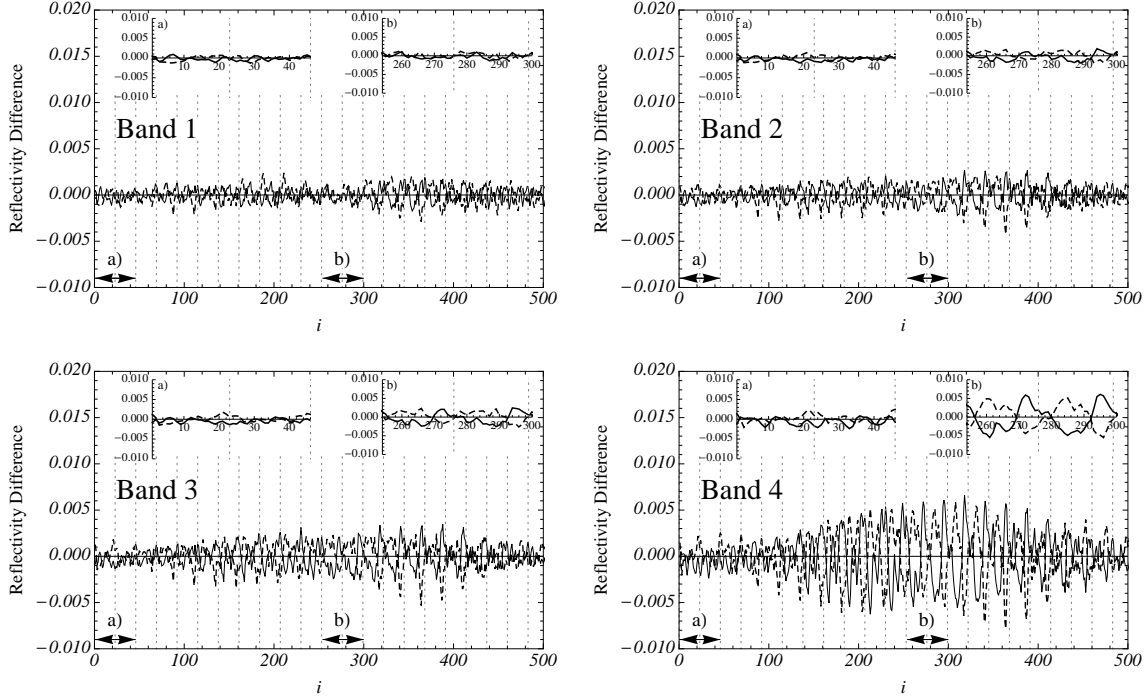


Figure 15. The difference between the light curve with the best-fit values ($\zeta_{\text{est}} = 44.4^\circ$, $\Theta_{S,\text{est}} = 179.8^\circ$), with a 3.3° shifted obliquity ($\zeta = 47.7^\circ$, $\Theta_S = 179.8^\circ$; solid curves) and with a -3.3° shifted obliquity ($\zeta = 41.1^\circ$, $\Theta_S = 179.8^\circ$; dashed curves)

The explicit equations of $\theta_V(\phi; \Phi; \zeta)$, $\theta_{I,+}(\phi; \Theta; \Phi; \zeta)$, and $\theta_{I,-}(\phi; \Theta; \Phi; \zeta)$ are expressed as

$$\theta_V(\phi; \Phi; \zeta) = \cos^{-1} \left(-\frac{\cos(\phi + \Phi) \sin \zeta}{\sqrt{\cos^2 \zeta + \cos^2(\phi + \Phi) \sin^2 \zeta}} \right), \quad (\text{A3})$$

$$\theta_{I,+}(\phi; \Theta; \Phi; \zeta) = \cos^{-1}(\eta(\phi + \Phi; \Theta; \zeta)) \quad (\text{A4})$$

$$\theta_{I,-}(\phi; \Theta; \Phi; \zeta) = \cos^{-1}(-\eta(\phi + \Phi; \Theta; \zeta)), \quad (\text{A5})$$

where

$$\eta(\phi; \Theta; \zeta) \equiv \left(\frac{\cos \zeta \cos \Theta \cos \phi + \sin \Theta \sin \phi}{\sqrt{\cos^2 \zeta \cos^2 \Theta \cos^2 \phi + 2 \cos \zeta \sin \Theta \cos \Theta \sin \phi \cos \phi + \sin^2 \zeta \cos^2 \Theta + \sin^2 \Theta \sin^2 \phi}} \right).$$

B. PIXELIZATION OF PLANETARY SURFACE

We note here that the design matrix alone does not fully specify \mathbf{m}_{est} without any constraints. We perform singular value decomposition (SVD) of the design matrix:

$$G = U \Lambda V^T, \quad (\text{B1})$$

where U and V are $N \times N$ and $M \times M$ unitary matrices, and Λ is a $N \times M$ matrix:

$$\Lambda = \begin{pmatrix} \tilde{\Lambda} \\ 0 \end{pmatrix}, \quad (\text{B2})$$

$$\tilde{\Lambda} \equiv \text{diag}(l_1, \dots, l_M), \quad (\text{B3})$$

with l_j being the singular value of G in descending order. We denote the j -th row of V by the vector \mathbf{v}_j .

In general, there are N_{null} zero components in the singular values, ($l_{M-N_{\text{null}}+1} = l_{M-N_{\text{null}}+2} = \dots = l_M = 0$). Corresponding \mathbf{v} vectors are called null vectors. Linear combination of the null vectors do not affect the data vector,

$$G \sum_{s=1}^{N_{\text{null}}} \beta^{\text{null},s} \mathbf{v}^{\text{null},s} = \mathbf{0},$$

$$\mathbf{v}^{\text{null},s} \equiv \mathbf{v}_{M-N_{\text{null}}+s}, \quad (\text{B4})$$

where $\beta^{\text{null},s}$ is an arbitrary coefficient. Equation (B4) indicates that the predicted model has a freedom to add any linear combination of the null vectors.

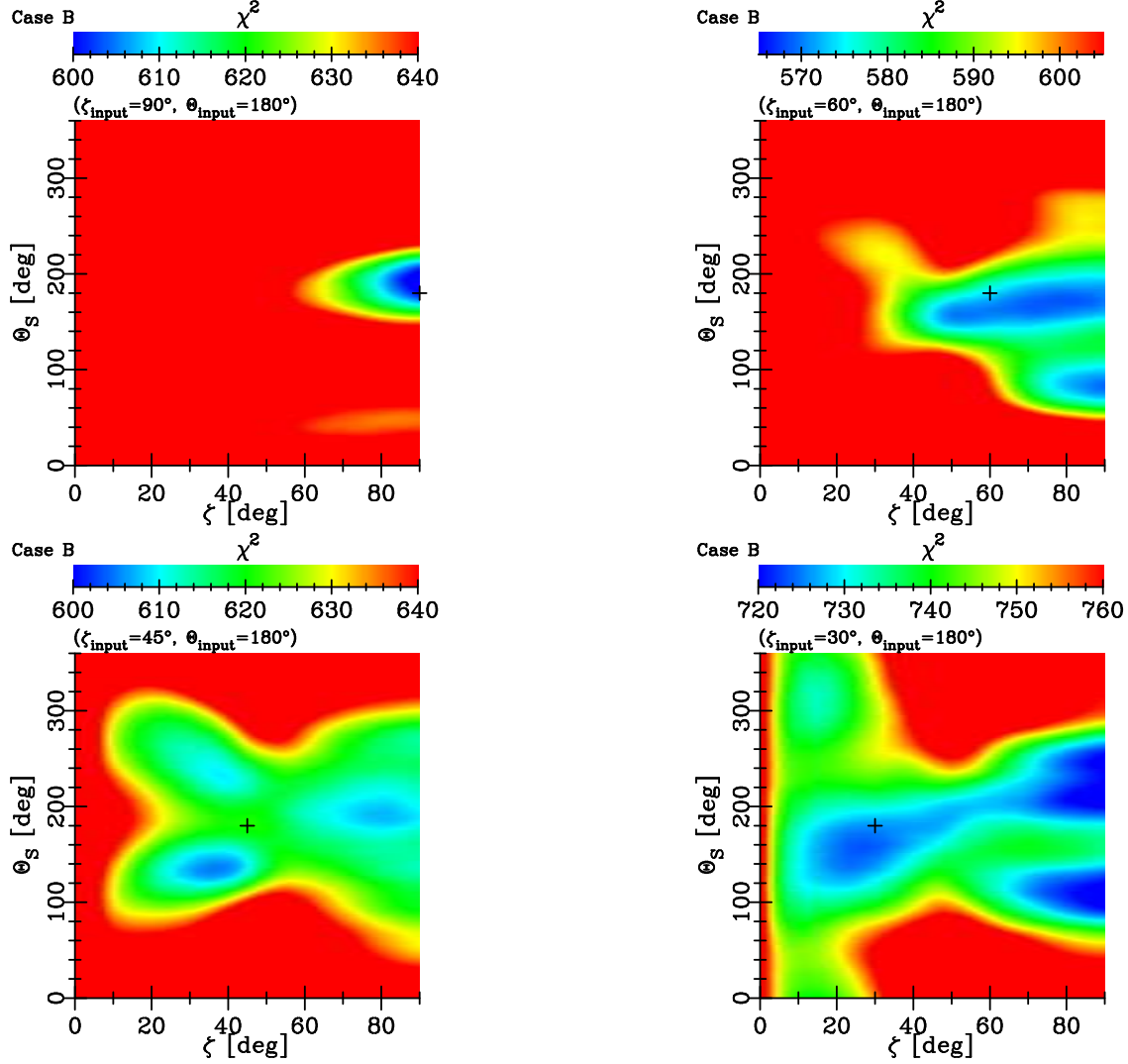


Figure 16. Same as Figure 14 for the case B.

Solid curve in the left panel of Figure 17 displays the singular values in descending order (this plot is referred to as *spectrum of data kernel*) in the case of $M_\phi = N_\Phi = N_\Theta = 23$ and $N_\theta = 9$. Due to the truncation errors, the singular values obtained by numerical algorithms of the SVD do not vanish exactly. However, only three of them ($j = M - 2, M - 1$, and M) have significantly smaller values than other diagonal elements. We regard the elements smaller than 10^{-4} as zero and the corresponding row vectors of V as null vectors.

Although the presence of null vectors does not directly lead to the indefiniteness if one adopts the inequality constraints (eq.[17]), smaller number of null vectors makes easier to interpret an uncertainty of the recovered map. Therefore it is useful to know how the number of null vectors depends on pixelization. Then one can decrease the number of the null vectors by an appropriate choice of pixelization. For simplicity, we fix $M_\phi = N_\Phi = N_\Theta$ and estimate the number of the null vectors by changing M_θ and M_ϕ . The left panel in Figure 17 shows the spectra of the singular value in descending order for three cases; $M_\phi = 22$ (dotted), $M_\phi = 23$ (solid) and $M_\phi = 24$ (dashed) with $M_\theta = 9$. The case of $M_\phi = 23$ has a steeper spectrum than the others. It means that the choice of $M_\phi = 23$ makes easier to identify null vectors. The right panel of Figure 17 indicates the number of null vectors for different sets of M_ϕ and M_θ , where we define vectors with singular value below 10^{-4} as null vectors. There is a general tendency that the odd number bin of M_ϕ has lesser N_{null} . Although the selection of smaller values of M_θ or M_ϕ generally leads to smaller number of the null vectors, it degrades the resolution of the map. As a compromise, we decided to adopt $M_\phi = 23$ and $M_\theta = 9$ for our fiducial values of the model in this paper. Figure 18 indicates the (unit) null vectors for $M_\phi = 23$ and $M_\theta = 9$, $\mathbf{v}^{\text{null},1} = \mathbf{v}_{M-2}$, $\mathbf{v}^{\text{null},2} = \mathbf{v}_{M-1}$, and $\mathbf{v}^{\text{null},3} = \mathbf{v}_M$ where \mathbf{v}_j is the j -th row of V . These three null vectors for $M_\phi = 23$ and $M_\theta = 9$ affect the large-scale structure of the map but do not change the small-scale distribution like the continents of the present Earth.

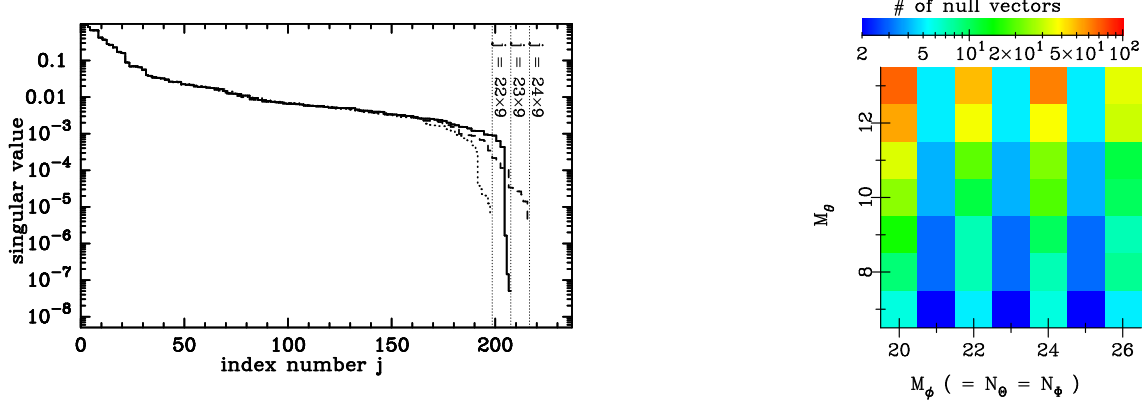


Figure 17. Singular values of the design matrix G in descending order (left panel). Dotted, solid, and dashed curves correspond to $M_\phi = 22$, $M_\phi = 23$ and $M_\phi = 24$, respectively. All curves assume $M_\theta = 9$. Right panel shows the number of vectors with singular values below 10^{-4} as a function of M_θ and $M_\phi = N_\theta = N_\phi$.

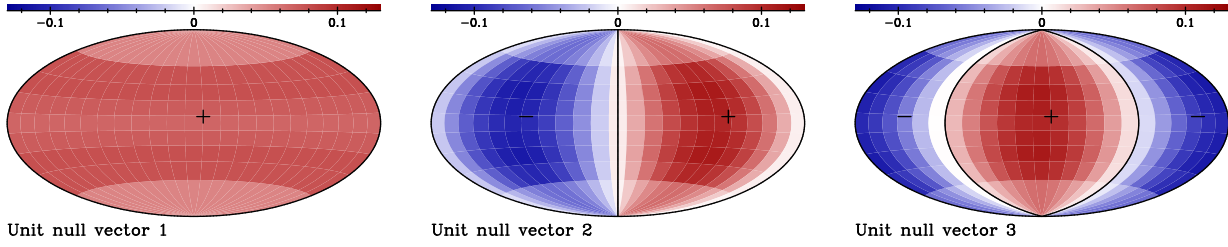


Figure 18. Three unit null vectors of our fiducial model $M_\phi = N_\phi = N_\theta = 23$ and $M_\theta = 9$. These maps are drawn using Hammer-Aitoff projection.

C. UNIQUENESS OF THE RECOVERED MAP

In this appendix, we examine whether there is the freedom for adding any components of null vectors to the recovered map derived by the BVLS. Because of the boundary condition for \mathbf{m}_{est} (eq. [17]), the estimated values of $m_{\text{est},j}$ are classified into 3 types —(a) $m_{\text{est},j} = 1$ (on the upper boundary), (b) $m_{\text{est},j} = 0$ (on the lower boundary), and (c) $0 < m_{\text{est},j} < 1$ (in between). We focus on types (a) and (b) to consider the uniqueness of our recovered map. We denote the index j of type (a) by $j-$ and that of type (b) by $j+$. In order not to violate the boundary condition, the linear combination of null vectors should satisfy the following constraints:

$$\sum_{k=1}^3 \beta^{\text{null},k} v_{j-}^{\text{null},k} \leq 0 \quad \text{for any } j-, \quad (\text{C1})$$

$$\sum_{k=1}^3 \beta^{\text{null},k} v_{j+}^{\text{null},k} \geq 0 \quad \text{for any } j+. \quad (\text{C2})$$

These inequalities can be solved in terms of $\beta^{\text{null},1}$ using the fact that v_j is positive (see Fig. [18]):

$$\beta^{\text{null},1} \leq -\frac{v_{j-}^{\text{null},2}}{v_{j-}^{\text{null},1}} \beta^{\text{null},2} - \frac{v_{j-}^{\text{null},3}}{v_{j-}^{\text{null},1}} \beta^{\text{null},3}, \quad (\text{C3})$$

$$\beta^{\text{null},1} \geq -\frac{v_{j+}^{\text{null},2}}{v_{j+}^{\text{null},1}} \beta^{\text{null},2} - \frac{v_{j+}^{\text{null},3}}{v_{j+}^{\text{null},1}} \beta^{\text{null},3}. \quad (\text{C4})$$

Then, the following inequality is required so that $\beta^{\text{null},1}$ exists,

$$\frac{v_{j-}^{\text{null},2}}{v_{j-}^{\text{null},1}} \beta^{\text{null},2} + \frac{v_{j-}^{\text{null},3}}{v_{j-}^{\text{null},1}} \beta^{\text{null},3} \leq \frac{v_{j+}^{\text{null},2}}{v_{j+}^{\text{null},1}} \beta^{\text{null},2} + \frac{v_{j+}^{\text{null},3}}{v_{j+}^{\text{null},1}} \beta^{\text{null},3}. \quad (\text{C5})$$

This equation can be reduced to the following expression:

$$\mathbf{p}_{(j+,j-)} \cdot \mathbf{b} \geq 0 \quad (\text{C6})$$

where $\mathbf{p}_{(j+,j-)}$ and \mathbf{b} are 2-dimensional vectors:

$$\mathbf{p}_{(j+,j-)} \equiv \left(\frac{v_{j+}^{\text{null},2}}{v_{j+}^{\text{null},1}} - \frac{v_{j-}^{\text{null},2}}{v_{j-}^{\text{null},1}}, \frac{v_{j+}^{\text{null},3}}{v_{j+}^{\text{null},1}} - \frac{v_{j-}^{\text{null},3}}{v_{j-}^{\text{null},1}} \right) \quad (\text{C7})$$

$$\mathbf{b} \equiv (\beta^{\text{null},2}, \beta^{\text{null},3}) \quad (\text{C8})$$

Now let us denote the arguments of \mathbf{p} and \mathbf{b} by $\theta_\alpha(j+,j-)$ and θ_β , respectively, i.e.,

$$\mathbf{p}_{(j+,j-)} \equiv |\mathbf{p}_{(j+,j-)}| (\cos \theta_\alpha(j+,j-), \sin \theta_\alpha(j+,j-)), \quad (\text{C9})$$

$$\mathbf{b} \equiv |\mathbf{b}| (\cos \theta_\beta, \sin \theta_\beta). \quad (\text{C10})$$

If $|\mathbf{b}| \neq 0$, the solution of equation (C6) should satisfy

$$\theta_\alpha(j+,j-) \leq \theta_\beta \leq \theta_\alpha(j+,j-). \quad (\text{C11})$$

We calculate the above constraints for any combination of $(j+,j-)$, and found that there are no θ_β that satisfies equation (C11) for all combinations of $(j+,j-)$. Thus, the only solution allowed is $\beta^{\text{null},2} = \beta^{\text{null},3} = 0$. From equations (C3) and (C4), it follows that $\beta^{\text{null},1} = 0$. As a result, we confirmed that there is no room for adding extra components of null vectors to the recovered map.

D. ANALYSIS WITH THE EXTENDED INFORMATION CRITERION

In this appendix, we try to perform the model selection of ζ_{est} and $\Theta_{S,\text{est}}$ using the extended information criterion (EIC; e.g. Efron 1983; Konishi & Kitagawa 1996; Ishiguro et al. 1997) based on the Kullback-Leibler divergence. Although the model selection for ζ and Θ_S based on the χ^2 minimization is very convenient, the result might be biased for large noises compared with signal. The EIC is a bootstrap-based extension of Akaike Information Criterion defined as

$$EIC \equiv -2 \sum_{i=1}^N \log \mathcal{F}(\mathbf{a} | \mathbf{m}_{\text{est}}, \zeta_{\text{est}}, \Theta_{S,\text{est}}) + 2b_B, \quad (\text{D1})$$

where $\mathcal{F}(\mathbf{a} | \mathbf{m}_{\text{est}}, \zeta_{\text{est}}, \Theta_{S,\text{est}})$ indicates the likelihood function. The best model or parameter set are given by minimizing the EIC. The bootstrap estimate of the bias b_B is given by (Ishiguro et al. 1997):

$$b_B = \frac{1}{N_{\text{EIC}}} \sum_{k=1}^{N_{\text{EIC}}} [\log \mathcal{F}(\mathbf{a}^{*,k} | \mathbf{m}_{\text{est}}^{*,k}, \zeta_{\text{est}}^{*,k}, \Theta_{S,\text{est}}^{*,k}) - \log \mathcal{F}(\mathbf{a}^{*,k} | \mathbf{m}_{\text{est}}, \zeta_{\text{est}}, \Theta_{S,\text{est}}) + \log \mathcal{F}(\mathbf{a} | \mathbf{m}_{\text{est}}, \zeta_{\text{est}}, \Theta_{S,\text{est}}) - \log \mathcal{F}(\mathbf{a} | \mathbf{m}_{\text{est}}^{*,k}, \zeta_{\text{est}}^{*,k}, \Theta_{S,\text{est}}^{*,k})], \quad (\text{D2})$$

where $\mathbf{a}^{*,k}$, $\mathbf{m}_{\text{est}}^{*,k}$, $\zeta_{\text{est}}^{*,k}$, and $\Theta_{S,\text{est}}^{*,k}$ indicate the k -th bootstrap resample, \mathbf{m}_{est} , ζ_{est} , and $\Theta_{S,\text{est}}$, and N_{EIC} is the number of trials of the bootstrap. Assuming a Gaussian distribution for \mathbf{a} with $\sigma = 1$, we derive

$$EIC = |\mathbf{a} - G(\zeta, \Theta_S) \mathbf{m}_{\text{est}}|^2 + N \log 2\pi + b_B$$

$$b_B = -\frac{1}{2N_{\text{EIC}}} \sum_{k=1}^{N_{\text{EIC}}} \left[\left| \mathbf{a}^{*,k} - G(\zeta^{*,k}, \Theta_S^{*,k}) \mathbf{m}_{\text{est}}^{*,k} \right|^2 - \left| \mathbf{a}^{*,k} - G(\zeta, \Theta_S) \mathbf{m}_{\text{est}} \right|^2 + \left| \mathbf{a} - G(\zeta, \Theta_S) \mathbf{m}_{\text{est}} \right|^2 - \left| \mathbf{a} - G(\zeta^{*,k}, \Theta_S^{*,k}) \mathbf{m}_{\text{est}}^{*,k} \right|^2 \right]. \quad (\text{D3})$$

We adopt $N_{\text{EIC}} = 500$ for bootstrap resampling. Figure 19 display the dependence of the EIC on ζ_{est} and $\Theta_{S,\text{est}}$. The minimum EIC estimate listed in Table 4 is derived by a stepwise search at one degree interval.

The minimum EIC estimate listed in Table 4 is derived by a stepwise search at one degree interval. For the case A, the EIC provides the almost same values as listed in Table 4. Figure 19 displays the EIC map for the case B. Although the estimated obliquities for $\zeta_{\text{input}} = 30^\circ, 45^\circ$ and 60° by the EIC do not agree well with the input value (Table 4), the bias of $\zeta_{\text{input}} = 60^\circ$ and 30° seems to be improved compared to that from χ^2 fitting.

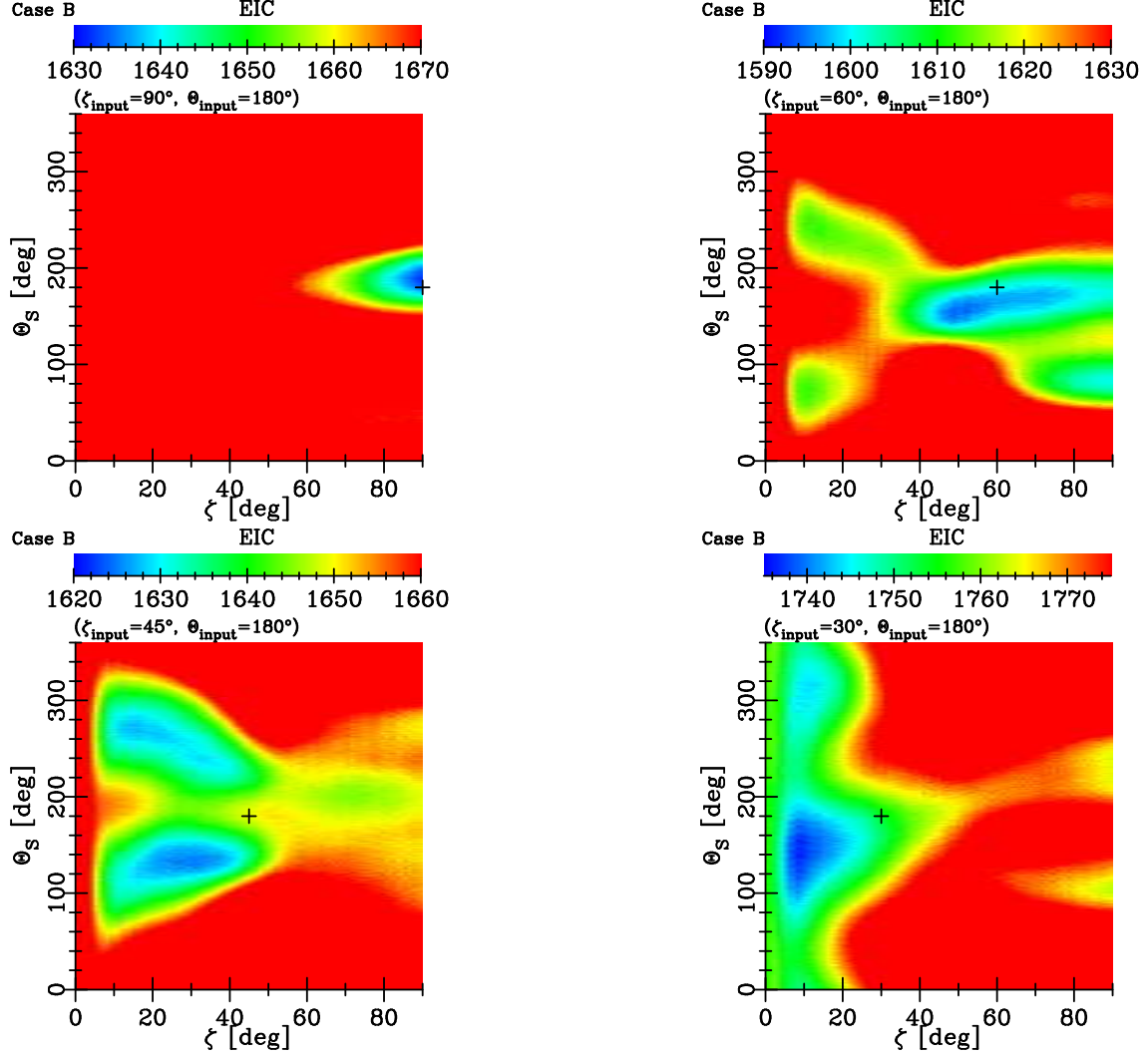


Figure 19. The ζ and Θ_S dependence of the EIC of the mock observational data set (case B) for $\Theta_{S,\text{input}} = 180^\circ$ and the different input obliquities, $\zeta_{\text{input}} = 90^\circ$ (top left), $\zeta_{\text{input}} = 60^\circ$ (top right), $\zeta_{\text{input}} = 45^\circ$ (bottom left), $\zeta_{\text{input}} = 30^\circ$ (bottom right). The input values are indicated by cross.

REFERENCES

- Agnor, C. B., Canup, R. M., & Levison, H. F. 1999, *Icarus*, 142, 219
- Arnold, L., Gillet, S., Lardi re, O., Riaud, P., & Schneider, J. 2002, *A&A*, 392, 231
- Chambers, J. E. 2001, *Icarus*, 152, 205
- Charbonneau, D., Berta, Z. K., Irwin, J., Burke, C. J., Nutzman, P., Buchhave, L. A., Lovis, C., Bonfils, X., Latham, D. W., Udry, S., Murray-Clay, R. A., Holman, M. J., Falco, E. E., Winn, J. N., Queloz, D., Pepe, F., Mayor, M., Delfosse, X., & Forveille, T. 2009, *Nature*, 462, 891
- Charbonneau, D., Brown, T. M., Noyes, R. W., & Gilliland, R. L. 2002, *ApJ*, 568, 377
- Colina, L., Bohlin, R. C., & Castelli, F. 1996, Absolute flux calibrated spectrum of Vega (Instrum. Sci. Rep. OSG-CAL-96-01, Baltimore: STSCI)
- Cowan, N. B., & Agol, E. 2008, *ApJ*, 678, L129
- Cowan, N. B., Agol, E., Meadows, V. S., Robinson, T., Livengood, T. A., Deming, D., Lisse, C. M., A’Hearn, M. F., Wellnitz, D. D., Seager, S., Charbonneau, D., & the EPOXI Team. 2009, *ApJ*, 700, 915
- Efron, B. 1983, *J. Amer. Statist. Assoc.*, 78, 316
- Ford, E. B., Seager, S., & Turner, E. L. 2001, *Nature*, 412, 885
- Fujii, Y., Kawahara, H., Suto, Y., Taruya, A., Fukuda, S., Nakajima, T., & Turner, E. L. 2010, *ApJ*, 715, 866
- Gaidos, E., & Williams, D. M. 2004, *New Astronomy*, 10, 67
- Ishiguro, M., Sakamoto, Y., & Kitagawa, G. 1997, *Ann. Inst. Statist. Math.*, 49, 411
- Kasdin, N. J., Spergel, D. N., Vanderbei, R. J., Cady, E., Savransky, D., Lisman, D., Shaklan, S., Lee, R., Egerman, R., Matthews, G., & Tenerelli, D. 2010, in *Bulletin of the American Astronomical Society*, Vol. 41, Bulletin of the American Astronomical Society, 287–+
- Knutson, H. A., Charbonneau, D., Allen, L. E., Fortney, J. J., Agol, E., Cowan, N. B., Showman, A. P., Cooper, C. S., & Megeath, S. T. 2007, *Nature*, 447, 183
- Kokubo, E., & Ida, S. 2007, *ApJ*, 671, 2082
- Konishi, S., & Kitagawa, G. 1996, *Biometrika*, 83, 875
- Lawson, C. L., & Hanson, R. J. 1974, *Solving least squares problems*
- . 1995, *Solving least squares problems (Revised edition)* (Society for Industrial and Applied Mathematics)

- Léger, A. et al. 2009, *A&A*, 506, 287
- Menke, W. 1989, *Geophysical data analysis: Discrete inverse theory*, ed. Menke, W.
- Montañés-Rodríguez, P., Pallé, E., Goode, P. R., & Martín-Torres, F. J. 2006, *ApJ*, 651, 544
- Nakajima, T., & Tanaka, M. 1983, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 29, 521
- Oakley, P. H. H., & Cash, W. 2009, *ApJ*, 700, 1428
- Pallé, E., Ford, E. B., Seager, S., Montañés-Rodríguez, P., & Vazquez, M. 2008, *ApJ*, 676, 1319
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical recipes in FORTRAN. The art of scientific computing*, ed. Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P.
- Russell, H. N. 1906, *ApJ*, 24, 1
- Savransky, D., Kasdin, N. J., & Cady, E. 2010, *PASP*, 122, 401
- Seager, S., Whitney, B. A., & Sasselov, D. D. 2000, *ApJ*, 540, 504
- Sudarsky, D., Burrows, A., Hubeny, I., & Li, A. 2005, *ApJ*, 627, 520
- Swain, M. R., Vasisht, G., & Tinetti, G. 2008, *Nature*, 452, 329
- Swain, M. R., Vasisht, G., Tinetti, G., Bouwman, J., Chen, P., Yung, Y., Deming, D., & Deroo, P. 2009, *ApJ*, 690, L114
- Tinetti, G., Meadows, V. S., Crisp, D., Fong, W., Fishbein, E., Turnbull, M., & Bibring, J. 2006a, *Astrobiology*, 6, 34
- Tinetti, G., Meadows, V. S., Crisp, D., Kiang, N. Y., Kahn, B. H., Fishbein, E., Velusamy, T., & Turnbull, M. 2006b, *Astrobiology*, 6, 881
- Tinetti, G., Vidal-Madjar, A., Liang, M., Beaulieu, J., Yung, Y., Carey, S., Barber, R. J., Tennyson, J., Ribas, I., Allard, N., Ballester, G. E., Sing, D. K., & Selsis, F. 2007, *Nature*, 448, 169
- Vidal-Madjar, A., Désert, J., Lecavelier des Etangs, A., Hébrard, G., Ballester, G. E., Ehrenreich, D., Ferlet, R., McConnell, J. C., Mayor, M., & Parkinson, C. D. 2004, *ApJ*, 604, L69
- Vidal-Madjar, A., Lecavelier des Etangs, A., Désert, J., Ballester, G. E., Ferlet, R., Hébrard, G., & Mayor, M. 2003, *Nature*, 422, 143
- Williams, D. M., & Kasting, J. F. 1997, *Icarus*, 129, 254
- Williams, D. M., & Pollard, D. 2003, *International Journal of Astrobiology*, 2, 1
- Williams, P. K. G., Charbonneau, D., Cooper, C. S., Showman, A. P., & Fortney, J. J. 2006, *ApJ*, 649, 1020
- Wolf, N. J., Smith, P. S., Traub, W. A., & Jucks, K. W. 2002, *ApJ*, 574, 430